Mesoscale wind fluctuations over Danish waters

Vincent, Claire Louise

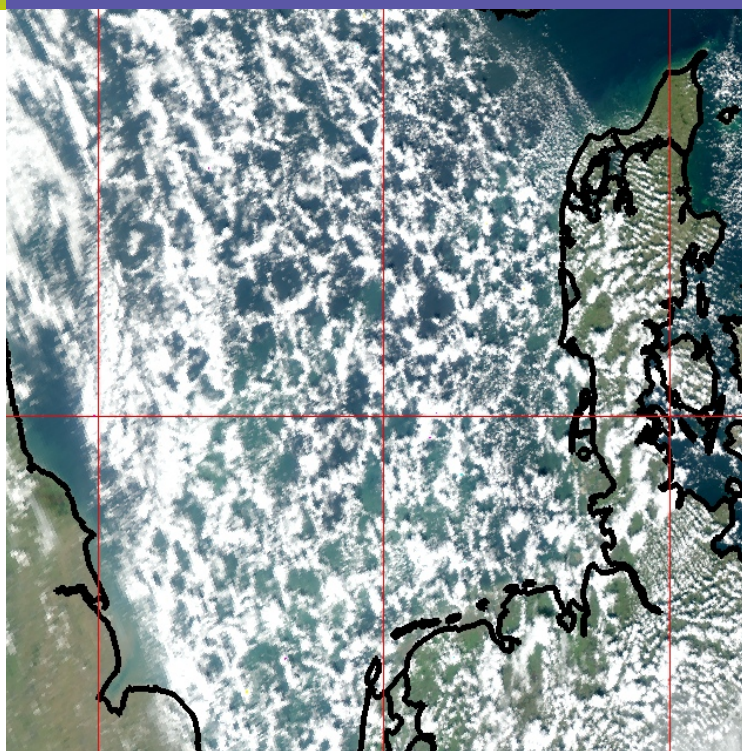
Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

Mesoscale wind fluctuations over Danish waters

Risø-PhD-Report

Claire Louise Vincent
Risø-PhD-70(EN)
December 2010



Risø DTU
National Laboratory for Sustainable Energy

Mesoscale wind fluctuations over Danish waters

Claire Louise Vincent

Roskilde 2010
Risø-PhD-70(EN)

Author: Claire Louise Vincent
Title: Mesoscale Wind Fluctuations over Danish Waters
Division: Wind Energy Division

Risø-PhD-70(EN)
December 2010

Abstract:

Mesoscale wind fluctuations affect the large scale integration of wind power because they undermine the day-ahead predictability of wind speed and power production, and because they can result in large fluctuations in power generation that must be balanced using reserve power. Large fluctuations in generated power are a particular problem for offshore wind farms because the typically high concentration of turbines within a limited geographical area means that fluctuations can be correlated across large numbers of turbines. Furthermore, organised mesoscale structures that often form over water, such as convective rolls and cellular convection, have length scales of tens of kilometers, and can cause large wind fluctuations on a time scale of around an hour.

This thesis is an exploration of the predictability of mesoscale wind fluctuations using observations from the world's first two large offshore wind farms - Horns Rev I in the North Sea, and Nysted in the Baltic Sea.

The thesis begins with a climatological analysis of wind fluctuations on time scales of 1–10 hours at the two sites. A novel method for calculating conditional climatologies of spectral information is proposed, based on binning and averaging the time axis of the Hilbert spectrum. Results reveal clear patterns between wind fluctuations and locally observed meteorological conditions. The analysis is expanded by classifying wind fluctuations on time scales of 1–3 hours according to synoptic patterns, satellite pictures and wind classes. Results indicate that cold air outbreaks and open cellular convection are a significant contributor to mesoscale wind variability at Horns Rev.

The predictability of mesoscale wind fluctuations is tested by implementing standard statistical models that relate local wind variability to parameters based on a large scale weather analysis. The models show some skill, but only achieve a 15% improvement on a persistence forecast. The possibility of explicitly modelling mesoscale fluctuations in a mesoscale model is then examined using the weather research and forecasting (WRF) model. A set of case studies demonstrate that realistic hour-scale wind fluctuations and open cellular convection patterns develop in WRF simulations with 2km horizontal grid spacing.

The atmospheric conditions during one of the case studies are then used to initialise a simplified version of the model that has no large scale weather forcing, topography or surface inhomogeneties. Using the simplified model, the sensitivity of the modelled open cellular convection to choices in model setup and to aspects of the environmental forcing are tested. Finally, the cell-scale kinetic energy budget of the modelled cells is calculated, and it is shown that the buoyancy and pressure balance terms are important for cell maintenance. It is explained that the representation of mesoscale convection in a mesoscale model is not only important to end users such as wind farm operators, but to the treatment of energy transport within the boundary layer.

The thesis is submitted to the Danish Technical University in partial fulfilment of the requirements for the PhD degree.

ISBN 978-87-550-3864-6

Sponsorship:
Mesoscale Variability and the variability of wind and production of offshore wind farms - PSO-7141
Safewind - EU project 213740
Radar@Sea - 2009-1-0226
HREnsembleHR - PSO-6382

Pages:173
References:143

Information Service Department
Risø National Laboratory for Sustainable Energy
Technical University of Denmark
P.O.Box 49
DK-4000 Roskilde
Denmark
Telephone +45 46774005
bibl@risoe.dtu.dk
Fax +45 46774013
www.risoe.dtu.dk

Dansk Resumé

Mesoskala vind-fluktuationer påvirker storskala integrering af vindenergi fordi de underminerer den daglige forudsigelighed af vind hastighed og produktionen af energi og fordi de kan resultere i store fluktuationer i energi produktionen der efterfølgende må erstattes af reserve energi. Store fluktuationer i genereret energi er et særligt problem for havvindmølle parker da den typiske høje koncentrationen af turbiner i et begrænset geografisk område betyder at fluktuationer er korrelerede henover et stort antal turbiner. Organiserede mesoskala strukturer der oftest dannes over vand, såsom konvektive ruller og konvektions celler, har længdeskalaer på titals-kilometer og kan skabe store vind-fluktuationer med tidsskalaer på omkring en time.

Denne thesis er en undersøgelse af forudsigeligheden af mesoskala vind-fluktuationer ved at bruge observationer fra verdens to første store havvindmølle parker – Horns Rev I i Nordsøen samt Nysted i Østersøen.

Afhandlingen begynder med en klimatologisk analyse af vind-fluktuationer med tidsskalaer på 1-10 timer ved de to placeringer. En ny metode til at beregne betingede klimatologier af spectral information er foreslået baseret på 'binning' og midling af tidsaksen af Hilbert spektrummet. Resultaterne afslører klare mønstre mellem vind-fluktuationer og lokalt observerede meteorologiske betingelser. Analysen er udvidet ved at klassificere vind-fluktuationer på tidsskalaer af 1-3 timer i henhold til synoptiske mønstre, satellit billeder og vindklasser. Resultaterne indikerer at koldlufts udbrud og åben celle konvektion væsentligt bidrager til variabiliteten af mesoskala vinden ved Horns Rev.

Forudsigeligheden af mesoskala vind-fluktuationer er testet ved at implementere standard statistiske modeller der relaterer den lokale vind variabilitet til parametre baseret på storskalaet vejr analyse. Modellerne har en vis formåen men opnår kun en 15% forbedring ved en persistens prognose. Muligheden for eksplicit at kunne modellere mesoskala fluktuationer i en mesoskala model er undersøgt ved brug af vejrprognose modellen (WRF). Studier af forskellige meteorologiske tilfælde demonstrerer at realistiske timelige vind-fluktuationer og åben celle konvektionsmønstre udvikles i WRF simuleringer med 2km's horisontal celle afstand.

De atmosfæriske betingelser under en af disse tilfælde er brugt til at initialisere en simplificeret version af modellen der ikke inkluderer storskalaet vejrforcering, topografi og overflade inhomogeniteter. Følsomheden af de modellerede åbne celle konvektioner overfor model opsætning og aspekter af miljømæssige forceringer er testet ved brug af denne simplificerede model. Endelig er det celle-skala kinetiske energibudget beregnet og det er påvist at opdrift og tryk balance leddene er vigtige for at opretholde celle strukturen. Det er yderligere forklaret, at mesoskala konvektion repræsenteret i en mesoskala model ikke kun er vigtigt for slut-brugere, såsom vindmølle park operatører, men også for energi transporten i grænselaget.

Acknowledgements

This project was made possible by all the wonderful people who helped and supported me in all sorts of ways during the past 3 years.

I'd first like to thank my supervisors, Andrea Hahmann, Pierre Pinson and Gregor Giebel for their ongoing support and assistance throughout the various stages of the project. I'm also grateful to Henrik Madsen for the help in the first stages of the project, to Yubao Liu for hosting me for 2 months at NCAR, and to Mark Kelly for lots of advice and interesting conversations. Thanks to everyone in the meteorology program at Risø-DTU for a supportive and interesting scientific environment – in particular, thanks to Jake Badger, Søren Larsen and Xioali Larsén for helpful advice. I'm grateful to Bill and Jeff from the Bureau of Meteorology in Australia, who gave me the initial inspiration to start a PhD. Thanks to the three examiners for their careful reading, excellent questions and useful suggestions.

Thanks to everyone who made life as a PhD student so fun and interesting! Thanks to Ferhat and Alfredo for being excellent office mates in the beginning, and for being the ones to answer the silly questions when I was new. Thanks to Caroline and Joakim for sharing the joys and frustrations of WRF. Special thanks to the 'Ubuntu team' from the MET program for moral and technical support. Thanks to Joakim for helping me learn Danish and for translating the abstract. And a big thank you to Caroline, Ioanna and Rozenn for the friendship and understanding. To Oliver and Eleni – thanks for the friendship and chances to get away from the computer and enjoy the natural world.

I'm very grateful for the constant support and encouragement of my family. To Jill and John – your visits have meant something really special. Thanks to Ky and Michael, who have always opened their home to me in London and visited Denmark regularly. And finally, to Jeremy – for patience, humour, perspective and everything else.

Thanks to Risø-DTU for supporting my PhD. The project was partly financially supported by the Danish PSO projects HREnsembleHR (contract number PSO-6382), Mesoscale Variability and the variability of wind and production of offshore wind farms (contract number PSO-7141), and Radar@Sea (contract number 2009-01-0226) and by the EU project Safewind (contract number 213740). I'm also grateful to the NCAR RAL Visitor program, which supported my two month visit to NCAR in 2009. Thanks to Dong Energy and Vattenfall for the generous permission to use observational data from Horns Rev I and Nysted wind farms.

Contents

Chapter 1

Introduction

The wind speed fluctuates on all time scales, from turbulence, which occurs on time scales of seconds to minutes, to mesoscale variability with periods of tens of minutes to hours, to diurnal, synoptic, annual and even longer time scale variability. Variability on all of these time scales influences the production of energy from the wind in different ways.

As the amount of installed wind energy production increases in Denmark and globally, managing the fluctuating nature of the wind speed will continue to present new challenges. Understanding, predicting and mitigating these effects will take a central place in the future large scale integration of wind energy. It is obvious that meteorology will play an increasingly important role not only in characterising the mean wind, but in characterising and predicting the statistical properties of the wind. Indeed, it is not only the future integration of wind energy that will demand advanced understanding and modelling of weather-related variability. Wave power and solar power will also contribute to the portfolio of future energy generation and will be subject to similar challenges relating to the natural variability of the atmosphere and the oceans. Therefore, understanding the fluctuations in the wind is extremely relevant both to the management and planning of current wind farms, and more broadly to future energy systems with much higher penetration of renewable energy.

1.1 Mesoscale wind fluctuations

Mesoscale fluctuations in the atmosphere have a length scale between a few kilometres and several hundred kilometres (Pielke sr., 2002), and this implies time scales between a few minutes and several hours. They are of particular interest because they fall within the so-called spectral gap region of the spectrum of atmospheric variability (Stull, 1988). According to the theory of the spectral gap, there is little variance between the diurnal and semi-diurnal frequencies and the turbulent frequencies. However, it is obvious that there are many atmospheric phenomena that do actually contribute to the variability of the wind in the spectral gap frequencies. For example, Stull (1988) mentions large cumulus clouds as contributors

to mesoscale variability, numerous studies have demonstrated that cellular convection and convective rolls generate mesoscale variability (Atkinson and Zhang, 1996; Agee et al., 1973; LeMone, 1976), Bakan (1985) calculated spectra of surface observations of wind speed and temperature during open cellular convection events and noticed a distinct spectral peak for periods of 60 minutes, Gjerstad et al. (1995) suggested that mesoscale cellular convection was responsible for the observed lack of spectral gap in wind speed and temperature time series from unstable conditions over the Norwegian Sea, and Larsén and Larsen (2010) have reported an observed case of gravity waves contributing to wind fluctuations of period 12 minutes in stable conditions over the Baltic Sea. These phenomena are by no means constant features of the atmosphere, but occur during special atmospheric conditions that support their development. Although the range of frequencies corresponding to mesoscale variability in the atmosphere does not have precise boundaries (for example, Orlanski (1975) defined mesoscale effects as spanning thunderstorms and urban effects with leng a length scales as small as 2 km to fronts with length scales of up to 2000 km), we are here interested in mesoscale effects that cause hour-scale fluctuations in the wind speed.

Mesoscale convection over the North Sea was the subject of an in-depth measurement campaign in 1981 referred to as the KonTur (Convection and Turbulence) experiment. During the campaign, open cellular convection and convective rolls were studied using data collected from aircraft, satellite and ground based instruments. As part of the KonTur experiment, Busack et al. (1985) analysed the frequency of open and closed cellular convection in the Norwegian Sea, just north of the North Sea, based on NOAA 6 satellite images. They found that open cellular convection, which consists of rings of cloud with clear centres, was observed on 78 days in 1981, while closed cellular convection, which consists of rings of clear air with cloudy centres, was observed on just 5 days. A similar result is presented in the forward to a special publication about the experiment, where the number of open cellular convection, closed cellular convection and convective rolls are determined from satellite pictures for the German Bight for the period October 1976 to September 1980. During the period, there were 10 occurrences of convective rolls, 74 cases of open cellular convection and 26 cases of closed cellular convection (Hinzpeter, 1985). Closer to the equator, cellular convection is also a dominating feature of the marine boundary layer. Painemal et al. (2010) found open cellular convection in 27.3% of satellite pictures over San Félix Island (26.5°S,80.0°W) in 2003, and closed cellular convection in 38% of satellite pictures. In a review of mesoscale cellular convection, Atkinson and Zhang (1996) collated observational evidence to show that convective rolls can be found with wavelengths between 1 km and 13.7 km, open cells with diameters between 2 km and 100 km, and closed cells with diameters of 1 to 80 km. Assuming that the mesoscale cellular convection is advected past a wind farm location with the mean wind, these length scales suggest that these processes are an excellent candidate for inducing hour-scale wind fluctuations at a fixed point.

Mesoscale fluctuations are also fascinating from a numerical modelling point of view because they lie just on the border of predictability by mesoscale models. As suggested by Skamarock (2004), features smaller than around 7 times the model's horizontal grid spacing suffer from spectral attenuation. This means that even for a relatively high resolution mesoscale model with a horizontal grid spacing of 2 km, the smallest features that can be properly resolved have a length scale of around 14 km. For typical mesoscale features such as cellular convection, convective rolls and mesoscale convective systems, which have length scales on the order of tens of kilometres, this means that it is by no means obvious that they should be simulated accurately. Nonetheless, many studies exist where such features have been successfully simulated using mesoscale models. For example, Wang and Feingold (2009a,b) used the Advanced Weather Research and Forecasting model (WRF) with a horizontal grid

spacing of 300 m to model pockets of open cellular convection within marine stratocumulus, and Richardson et al. (2007) used WRF with a similar horizontal grid spacing to model open cellular convection on Mars. For stable boundary layer conditions, Larsén and Larsen (2010) successfully simulated an observed case of gravity waves over the Baltic Sea using WRF and showed that the shape of the mesoscale part of the u and v wind spectra were simulated correctly both in the presence of gravity wave fluctuations of period about 12 minutes, and in the presence of a convective boundary layer.

Simulating mesoscale fluctuations using large eddy simulation (LES) models, which have a higher resolution than mesoscale models and explicitly resolve the larger turbulent eddies, is also challenging. The domain must be large enough to contain a representative sample of the mesoscale phenomena, while having a horizontal grid spacing small enough that the spectral cut-off of the model falls within the inertial sub-range. These dual requirements mean that LES simulations of mesoscale phenomena demand computing resources that are presently prohibitive for operational applications. Nonetheless, in research applications, several studies have simulated process relevant to mesoscale wind fluctuations using LES modelling. For example, Müller and Chlond (1996) initialised a convective boundary layer using an LES model with a horizontal grid spacing of 100 metres and a domain size of 6.4 $\times$ 6.4 km. As the length scale of the dominant convective structures grew, they successively moved the simulation to a new grid with an increasingly larger horizontal grid spacing and domain size, using the previous simulation to periodically tile the larger scale simulation. The technique was successful for simulating open cellular convection, and it was found that the length scales of all fluctuating fields were strongly dominated by the size of the cells. In contrast, in a more recent experiment, Schröter et al. (2005) did away with the need for regridding and simulated cellular convection directly on an LES domain. They found that while the length scales of scalars and dynamic variables were dominated by the size of the cells, the spectrum of vertical velocity had the highest contribution to the energy at smaller scales. These results were strengthened by the further LES simulations of de Roode et al. (2004), who found that in an LES simulation of the convective boundary layer, the vertical velocity and virtual potential temperature were dominated by the length scale of the planetary boundary layer height, while the potential temperature and specific humidity fields were dominated by larger length scales.

These LES modelling efforts, together with observational evidence, demonstrate that the structure of mesoscale convection is more complex than the simple convection cells that can be simulated by a mesoscale model. However, LES models are presently not ready for operational use, particularly in the field of real-time weather forecasting, and it is therefore still important to understand the behaviour of mesoscale modelling in this context. Running mesoscale models is also relatively cheap and fast compared to LES modelling, and this makes it possible to perform a large number of experimental simulations. Furthermore, boundary layer convection has an important role in the transport of energy from the surface out of the boundary layer, so it is highly relevant to study the behaviour of mesoscale models under the dynamically active situations of cellular convection.

1.2 The impact of wind variability on power production

All wind fluctuations influence wind power production. For example, the intensity of turbulent fluctuations has an important effect on the structural loading of the wind turbine (eg. Bierbooms (2005); Sørensen et al. (2008a)) and must be considered in very short term turbine control (eg Kusiak et al. (2010)). Even very slow changes on seasonal, annual or even longer time scales are important, since they can have strong impact on the annual energy production of a wind farm (Pryor and Barthelmie, 2010).

Over a large spatial area containing many wind turbines, the turbulent fluctuations in the power produced by each turbine are expected to be smoothed (Sørensen et al., 2008b). On the other hand, mesoscale fluctuations, which occur on time scales of tens of minutes to hours, arise from meteorological phenomena that have length scales of several kilometers. This means that there can be a high spatial correlation between the wind speed at neighbouring turbines, and the fluctuations will not necessarily be smoothed in the same way as turbulent fluctuations. For onshore wind turbines, several studies have remarked on the benefit of spatial smoothing of aggregated power output due to a large geographical distribution of wind turbines (Landberg, 1997; Giebel, 2000; Focken et al., 2002). However, for offshore wind farms, where there is usually a high concentration of turbines within a limited geographical area, mesoscale fluctuations may translate into large fluctuations in the aggregated power output from a whole wind farm, or even groups of wind farms (Akhmatov et al., 2007). Mesoscale fluctuations therefore affect the real-time balance of the power system, since any rises and falls in wind power production must be compensated with reserve power.

An episode of large mesoscale wind fluctuations that was observed at the Horns Rev wind farm is shown in figure 1.1. The first panel shows the measured wind speed from an anemometer mounted at a height of 62 m on a meteorological mast near the wind farm, where the measurement resolution was ten minutes. The second panel shows the measured wind speed at the nacelle of each of the turbines in the wind farm. The wind direction was southerly for the first day of the event, then swings around to the west-southwesterly directions during the two periods of enhanced wind variability on days 18 and 19. Although the wind turbines at Horns Rev cover a spatial area of about 20 km^2, the fluctuations are clearly not significantly offset or uncorrelated from one another over the extent of the wind farm. Even the wake effect of the wind farm, which is well known to cause deficit in the mean wind speed (eg. Barthelmie et al. (2009), Lange et al. (2003)), has apparently not disrupted the highly correlated wind speed fluctuations amongst the turbines. This is translated directly into the large fluctuations in the aggregated power output from the whole wind farm which is shown (for the 68 out of 80 turbines that were reporting power during the event) in the bottom panel.

In Denmark, electricity producers, including wind farm operators, are required to submit day-ahead bids for their expected energy production. In the case of wind power, this day-ahead prediction of wind power is reliant on a good prediction of the short term evolution of the wind conditions. Errors in the prediction contribute to the overall imbalance in the system. The system imbalance (resulting from errors in prediction of demand as well as errors in prediction of supply) is continually adjusted on the balancing market, where electricity is bought and sold on an hourly basis by the Transmission System Operator (TSO), Energinet.dk. In other countries, similar systems exist with subtle variations in

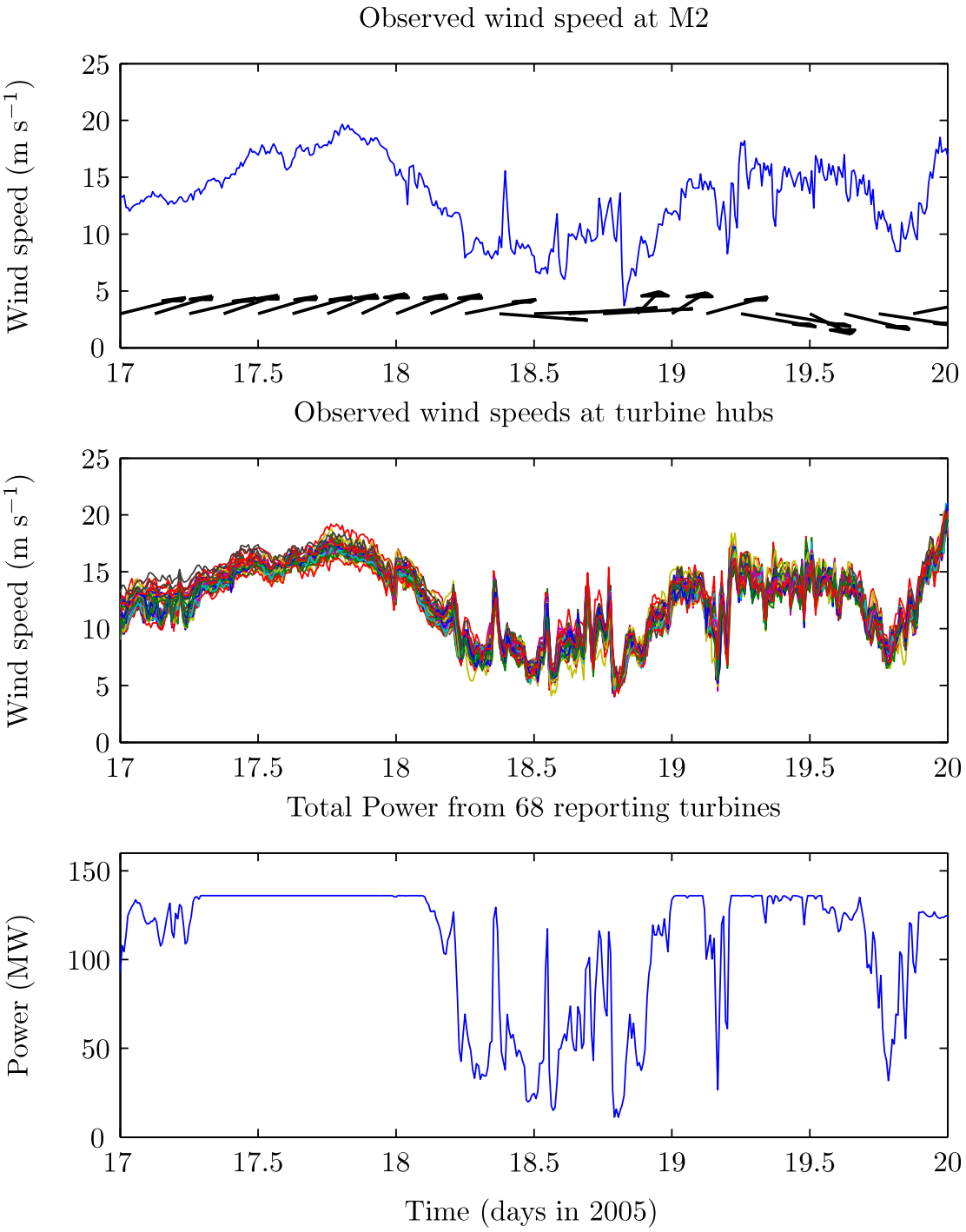


Figure 1.1: A three day period in 2005 when large amplitude mesoscale wind fluctuations were observed at the Horns Rev wind farm. Top panel: Wind speed and direction observed at a meteorological mast located to the northwest of the wind farm, at a height of 62 m above sea level. Middle panel: Wind speeds observed on the nacelle of each of the turbines in the wind farm, at a height of about 80 m above sea level. Lower panel: Aggregated power output for the wind farm, based on the 68 turbines that were reporting at the time.

the timing and frequencies of bidding and forecasting, and in the financial mechanisms for keeping the overall supply and demand of the electricity system in balance (Ackermann et al., 2007).

Electricity producers, and in particular wind farm operators, are penalised for day-ahead prediction errors that contribute to the imbalance in the system (Holttinen, 2005). That is, if the electricity system is up-regulating, and the wind energy production exceeds the predicted value, or if the electricity system is down-regulating and wind energy production fails to meet the predicted value, then financial penalties are applied. On the other hand, if the day-ahead prediction errors actually alleviate imbalance in the overall electricity system, then equivalent financial rewards are not applied. Therefore, there is a strong financial disincentive to make large errors in day-ahead wind power production forecasts. Mesoscale wind fluctuations, which have a time scale of minutes to hours, can contribute significantly to forecasting errors, since forecasts originating from large scale weather models will not contain any such fluctuations, and because the phase of individual fluctuations is difficult to predict.

The challenges posed by mesoscale wind fluctuations are not, however, just an artificial problem posed by the structure of the Danish electricity market. Ultimately, the supply of electricity must equal the demand at all times, and market mechanisms are usually designed to support the technical requirements of meeting the electricity demand. Technical requirements include having enough reserve power in the electricity system to meet any shortfall due to unexpected drops in wind power production, and having the right mix of generation capacity in the system to respond to changes in wind power production on a short time scale. For example, gas fired power stations and hydro-electric power stations can respond quickly to changes in demand, while coal fired power stations take several hours to respond to a sudden requirement for extra production (Tuohy et al., 2009). This means that the system must have enough fast-responding reserve power to counteract the effect of mesoscale wind fluctuations.

Many of the financial and technical challenges of managing the fluctuating nature of the wind can be at least partly alleviated by improvements in forecasting. Although day ahead, deterministic forecasts of the exact timing of hour-scale wind fluctuations are probably not possible using current technology and physical understanding, forecasts of the occurrence and amplitude of the phenomena could go a long way towards ensuring that the right type and amounts of reserve were available. The ability to predict day-ahead amplitude and occurrence of large wind fluctuations could actually contribute to a more general shift from deterministic forecasting and towards stochastic and probabilistic forecasting.

Fluctuations in the power produced by wind farms is closely coupled with the fluctuations in the wind, but the transfer function between wind and power fluctuations is complex and non-linear due to the non-linear power curve of the turbines, the smoothing effect of turbines that see the same fluctuation at a slightly different time, and the spatial correlation structure of the wind fluctuations. The precise way in which wind fluctuations are transformed into power fluctuations is an interesting topic in itself, but it is not addressed in this thesis. The analysis is restricted to the fluctuations in the wind itself.

1.3 Predictability and forecasting

During the past two decades, there has been a focus on the development of forecasting models for wind power prediction (Giebel, 2011). Models span timescales from a few minutes ahead to day ahead or even longer. The shorter timescales, from minutes to hours, are relevant to

the balancing and scheduling of power systems that include significant levels of wind power. Such forecasts are usually based on statistical models such as ARIMA models or neural networks, often with adaptive estimation of coefficients and various online measurements as inputs (Giebel, 2011; Costa et al., 2008). These models are not the focus of this thesis, since we are concerned with wind fluctuations that have periods longer than the time frames of these models themselves.

The role of day-ahead wind and power forecasting is to guide the day-ahead bids of expected wind power production on the electricity spot market. Day ahead wind power forecasting systems are usually combinations numerical weather prediction models and physical and / or statistical downscaling. For example, the *wind power prediction tool* (WPPT), first initiated in 1992 for forecasting wind power production in Denmark, is a statistical model that combines measurements of wind speed and wind power with information from meteorological models to forecast wind power for reference wind farms, that are then up-scaled to provide forecasts for areas of interest (Nielsen, 1999). In contrast, *Prediktor* is a physical model that uses input from a numerical weather prediction model, and adds sub-grid-scale local effects such as roughness changes, speed-up over hills and shelter from obstacles using the Risø-DTU software WAsP. Statistical corrections are finally applied to adjust for systematic errors (Landberg, 1999, 2001).

As argued by Pinson et al. (2007), the market value of wind power can be enhanced by incorporating probabilistic information into the decision making processes for optimal bidding strategies. For example, the WEPROG forecasting system (Möhrlen, 2004) uses a 75 member ensemble of mesoscale models to estimate uncertainties in wind power forecasts and WPPT includes estimates of forecast uncertainty using statistical methods. Approaching the problem from a different angle, Potter et al. (2009) propose a forecasting tool for *rapid ramp events*, which are defined as changes of more than 10% of the nameplate capacity of the wind farm within an hour, and Lange et al. (2010) argue that normal ways of combining and optimising forecasts tend to wash out ramps, and they propose a ramp forecasting system that predicts the timing, magnitude, ramp rate and duration of ramps in wind speed.

Mesoscale wind variability contributes to the uncertainty in the day-ahead wind speed, but it can also be considered as a stochastic process imposed on a slowly varying background wind, which itself does not necessarily have low predictability. For example, the spatially averaged wind conditions behind a front could be known with reasonable certainty based on output from a numerical weather prediction ensemble model, but it might also be known that such conditions were likely to lead to mesoscale fluctuations, the phase of which is impossible to forecast on a day-ahead time scale. In this way, the mesoscale wind fluctuations can be considered as decoupled from the slow fluctuations in the mean wind, and the problem of day-ahead forecasting of mesoscale wind fluctuations can be reduced to the task of forecasting their frequency and amplitude but not their phase. The interplay between variability and predictability is in itself an interesting and important area for future research.

1.4 The predictability of mesoscale scale wind fluctuations over Danish waters

Mesoscale wind fluctuations are fascinating and challenging not only because they have a vital importance and relevance to the large scale integration of wind energy, but because they present new and unresolved issues in meteorology and mesoscale modelling. They sit between the established theories of turbulence, which is often studied by considering short periods of high resolution observations as statistically stationary and homogeneous, and large scale meteorology, where hour-scale wind fluctuations are generally ignored. Therefore, the toolbox for studying mesoscale phenomena, particularly in a climatological sense, is not as well established as the toolboxes for studying turbulence or the mean flow.

This thesis is an attempt to contribute to the understanding of mesoscale variability and its predictability, with regard to meteorology, mesoscale modelling and methodological development for statistical analysis. Particular reference is made to the way that these challenges are framed in the context of wind power systems. The idea of mesoscale wind fluctuations is not a strictly defined topic that sits cleanly within well delineated boundaries or within a particular established branch of atmospheric science. Therefore, the work draws on numerous theoretical ideas from statistics, turbulence, synoptic and mesoscale meteorology and numerical modelling. The work is arranged in nine sections relating to the statistical analysis of wind fluctuations, the classification of wind fluctuations according to large scale meteorology and the modelling of open cellular convection using the WRF model. Three papers were written during the course of the PhD, and these are included in the appendix.

As is common practice in studies of meteorological phenomena, the study of mesoscale wind fluctuations begins with an analysis of its climatology. However, the concept of a climatology of fluctuations requires a bridge between the mean conditions and the instantaneous fluctuations. A Fourier spectrum expresses the average weights of the various periodicities present in a time series. However, studying the time dependent evolution of wind fluctuations requires a methodology for calculating the instantaneous frequencies in the time series, and for this the Hilbert-Huang transform (Huang and Wu, 2008) was shown to be an ideal tool. The characteristics of wind fluctuations do not only depend on time, but also on the meteorological conditions such as wind direction or air-sea temperature difference. Therefore, new methodological procedures are suggested for calculating a conditional climatology of wind fluctuations based on the Hilbert-Huang transform. The description of the methodology can be found in chapter 3 and in the paper Vincent et al. (2010a).

The suggested methodology is applied to observational data from the first two large offshore wind farms in the world, Nysted (in the Baltic Sea) and Horns Rev I (in the North Sea). Climatologies of wind variability are presented as conditional spectra, where the time evolving axis of the spectrum is binned and averaged to suggest the average wind variability as a function of wind direction, wind speed, pressure tendency, air-sea temperature difference, and time of year. The results are presented in chapter 4 and in the paper Vincent et al. (2010b). These results are specific to the two sites studied, but certain aspects of the results may point to more general conclusions, such as the relationship between variability and air-sea temperature difference. To establish the generality of other results, the analysis of further sites would be required. For example, the difference in wind variability in flow from the sea compared to in flow from the land at Horns Rev I is probably a function both of the different meteorological conditions that exist over the sea and of the fact that many

dynamic weather systems approach the wind farm from the westerly direction. Unravelling the relative impacts of these two effects could be achieved by studying similar sites on the east coasts of continents, such as at the sites of large offshore wind farms planned off the English North Sea coast, or the east coast of the United States.

The results from the conditional climatology of wind fluctuations at Horns Rev I and Nysted suggest fascinating clues about the physical phenomena that could be contributing the wind variability, but like all purely statistical analysis, further evidence from physical reasoning and observational evidence was required. Variability, which is defined using an index based on the adaptive spectral analysis of time series at Horns Rev, is studied in relation to visual satellite pictures and data from a large scale weather model. It is shown that there are strongly favoured synoptic patterns for the development of mesoscale wind fluctuations, and that open cellular convection can be observed over the North Sea for many of the observed cases of large amplitude mesoscale wind fluctuations. These results are collected in chapter 5. Since open cellular convection is identified as an important risk factor for large amplitude wind variability, a literature review of mesoscale cellular convection in general is presented in chapter 6.

The conditional climatology and analysis of satellite pictures and synoptic types provide a strong case that there should be some skill in predicting mesoscale wind variability based on direct and derived variables from the large scale flow. This hypothesis is tested, and several statistical models are implemented in an attempt to predict mesoscale wind variability using input from the ECMWF Interim Reanalysis product (Simmons et al., 2006/2007). The decisions about the inputs to the models draw strongly on the results that are established in the previous chapters, and include derived parameters such as the synoptic category and the air-sea temperature difference. The purpose of this modelling exercise is not to develop new or particularly sophisticated statistical models, but to demonstrate the level of skill in predicting mesoscale variability based on large scale models that do not contain any mesoscale variability themselves. Linear, auto-regressive models and models which took into account the non-linear relationship between the variability and the various predictors are tested. The results of the statistical models are found in chapter 7.

The statistical models showed some skill, but could only achieve about a 15% improvement on a persistence model. A totally different approach is to downscale the large scale weather conditions using a mesoscale model. In this case, the large scale weather model (or analysis) provides the boundary conditions and initial conditions to a series of nested limited area models, which should be capable of simulating smaller and smaller scales of atmospheric motion. As already discussed, mesoscale phenomena such as open cellular convection lie on the border of predictability by mesoscale modelling, and it was therefore important to test the outcome of simulating mesoscale variability on a phenomenological case study basis. Eight cases of observed severe wind variability observed at the Horns Rev wind farm between 2000 and 2003 are simulated using the WRF model, together with a control case where the mesoscale variability was extremely low. It was by no means obvious that the mesoscale model, which was run at a horizontal resolution of 2 km, could simulate the events, but the results, presented in chapter 8, are in fact promising. All cases appear to develop cellular convection patterns with length scales of the same order of magnitude as those observed in the observed satellite pictures for the events, and the fluctuations seen in time series of wind speed near the Horns Rev wind farm are realistic.

After showing that it was possible to simulate cellular convection and mesoscale wind fluc-

tuations with the WRF model, the performance of the model is scrutinised in some detail by running a set of experiments using a simplified version of the model, where vertical profiles of potential temperature, vapour mixing ratio and wind speed are used to initialise a 3 dimensional simulation of wind blowing over a water surface. The simplified model contains no topography or large scale weather forcing, and has periodic boundary conditions in both the meridional and zonal directions.

WRF is a modular system with many options for various aspects of the physics and dynamics in the model. The simplified model provided an ideal modelling environment to test the setup of the model, which is relevant to the practical use of mesoscale models for forecasting mesoscale wind fluctuations, and to the understanding of the model itself. Experiments that tested the various diffusion options, the horizontal grid spacing and the boundary layer scheme are used to come up with a preferred model set-up. However, although there were some obvious criteria upon which to base the performance of the model (such as the robustness of the length scale of the cells to changing the horizontal grid spacing), there was no benchmark which could definitively reveal the 'best' or 'most realistic' simulation.

Having evaluated the setup of the idealised model, a further series of experiments which tested the sensitivity of the simulations to changes in the physical forcing are conducted. In interpreting these results, the important difference between analysing the physical response of the mesoscale model to certain environmental conditions, and analysing the physical response of the real atmosphere to the same environmental conditions must be respected. The experiments actually only reveal the physical response of the model, which gives clues but not definitive answers about the behaviour of the real atmosphere. Nonetheless, results such as the relationship between the cell diameter and the boundary layer height seem to be consistent with established observational evidence, even if they only cover a small fraction of the possible parameter space. The description of the model, and results from modelling experiments are found in chapter 9.

The energetics of the simulated open cellular convection is studied in further detail in chapter 10 and in the paper Vincent et al. (2011) by calculating the cell-scale kinetic energy budget of the modelled cells, drawing closely on the definitive work of LeMone (1973, 1976). In calculating such a budget, it is necessary to take the unresolved scales of variability into account, since the correlations between the turbulent fluctuations (which form the Reynolds stress terms in the Navier-Stokes equations) contribute to the budget, theoretically as the 'interscale transfer' that is often considered in the context of LES modelling. In the case of the mesoscale model, this term is responsible for all the energy that is ultimately removed from the model as a result of explicit and implicit diffusion. By taking the diffusion from the boundary layer scheme, the explicit horizontal diffusion and the numerical diffusion into consideration, it is shown that it is possible to balance the budget to a good approximation. To the best of my knowledge, this is the first time that all the terms in the cell-scale kinetic energy balance of open cellular convection have been calculated. Although one must be careful in drawing direct conclusions about the behaviour of the real atmosphere based on mesoscale modelling results, it seems likely that the qualitative features of the budget are realistic. Furthermore, the calculations provide a fascinating way of quantifying the relative contribution of the explicit diffusion to the energy dissipation in the model. The key result from the budget was that the buoyancy term was balanced by the pressure balance term, and that shear did not play an important role in the process.

To study mesoscale wind fluctuations within a solid statistical and physical framework re-

quires a interdisciplinary approach. From some perspectives, the work draws on ideas and methods from turbulence studies - for example in the partitioning of energy that was necessary to calculate the cell-scale kinetic energy budget. On the other hand, mesoscale fluctuations are different to turbulent fluctuations because they rarely appear in sufficiently long segments of time series for calculation of Fourier spectra based on statistically stationary conditions. This work therefore called for diverse analysis techniques. Where necessary, new techniques have been developed - for example, in the development of the conditional climatology of wind variability based on the Hilbert-Huang transform. In other areas, existing methodology has been applied to the problem, such as in the use of self-organising maps for synoptic classification, or in the use of the Weather Research and Forecasting model for mesoscale modelling.

Chapter 2

Site description and data

2.1 Introduction

The analysis of mesoscale wind variability that is presented in this thesis is primarily based on time series observations from meteorological masts at the Horns Rev I wind farm in the North Sea, and the Nysted wind farm in the Baltic Sea. The purpose of this chapter is to describe the locations, the measurement systems and the average wind conditions for the two sites.

2.2 Site Descriptions

The Horns Rev wind farm is situated in the North Sea, near the west coast of Denmark. It is shown in figure 2.1, where each dot within the wind farm area corresponds to one turbine. Most of the measurements of wind speed and meteorological conditions at the Horns Rev wind farm used in this thesis were measured at the M2 measurement mast, which is indicated to the northwest of the wind farm. When the wind blows from the southeasterly direction, the flow at M2 is affected by the wake of the wind farm. The approximate extent of the wake affected flow is shown in the upper panel of figure 2.1. Wind from the west has a long ocean fetch, extending as far as England in the southwest to northwesterly directions, and extending deep into the North Sea and the Norwegian Sea in the northwest to north-northwesterly direction. Flow from the north passes over complex topography in Norway before reaching the Horns Rev wind farm.

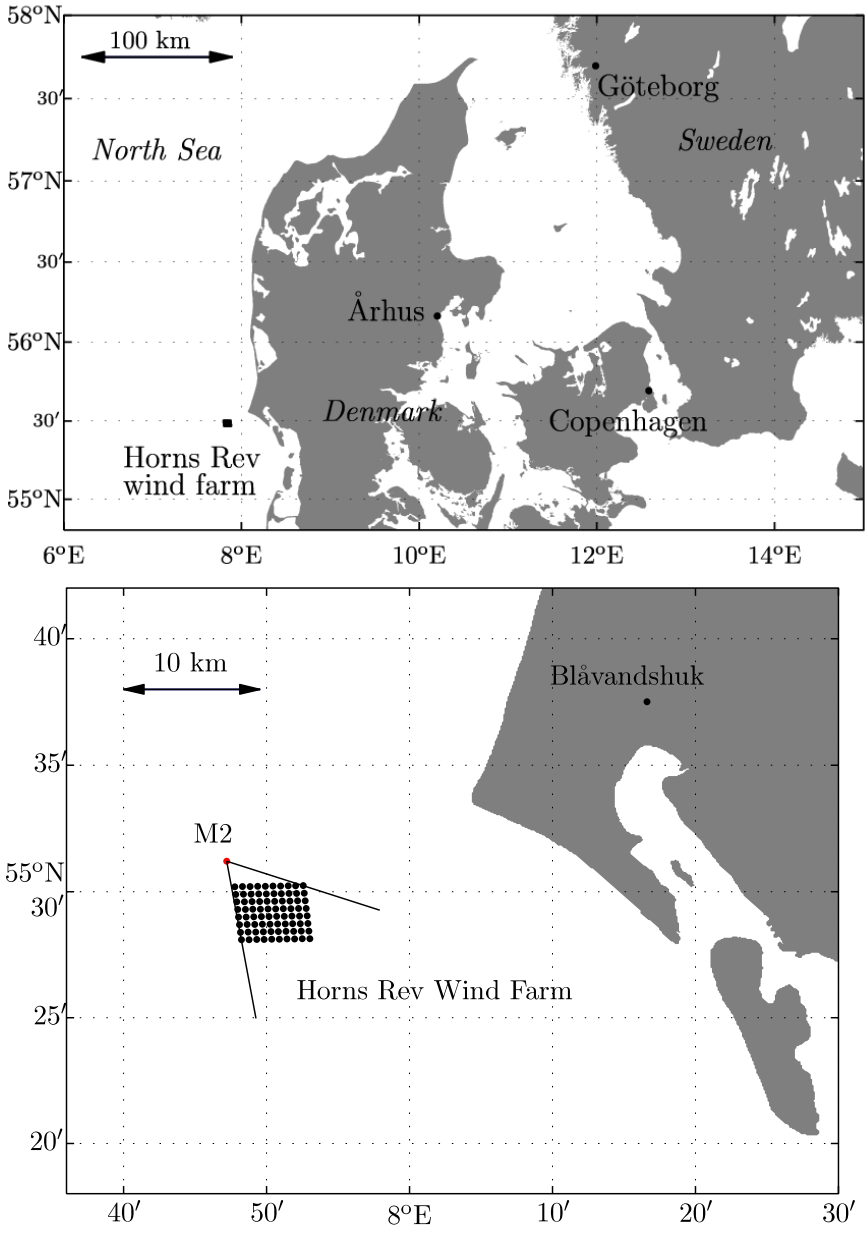


Figure 2.1: The layout and location of the Horns Rev wind farm with respect to the nearby coastline. Each black dot within the wind farm area indicates one turbine. The meteorological mast is shown as a red dot, labelled M2. The wind direction sector where the mast is under the influence of the wind farm wake is shown in the upper plot.

The Nysted wind farm is situated in the Baltic Sea, between Denmark and Germany. The location of the Nysted wind farm is shown in figure 2.2. The land surrounding the Nysted wind farm is more complex than the case of Horns Rev. The longest ocean fetch is in the northeast to easterly direction. Two meteorological masts were used for the analysis. The first is located at the southwest corner of the wind farm, very close to one of the turbines. For this mast, flow from the north to east-northeasterly directions are wake affected. The second mast is located to the east of the wind farm, so that flow from the southwest to northwesterly directions is wake affected.

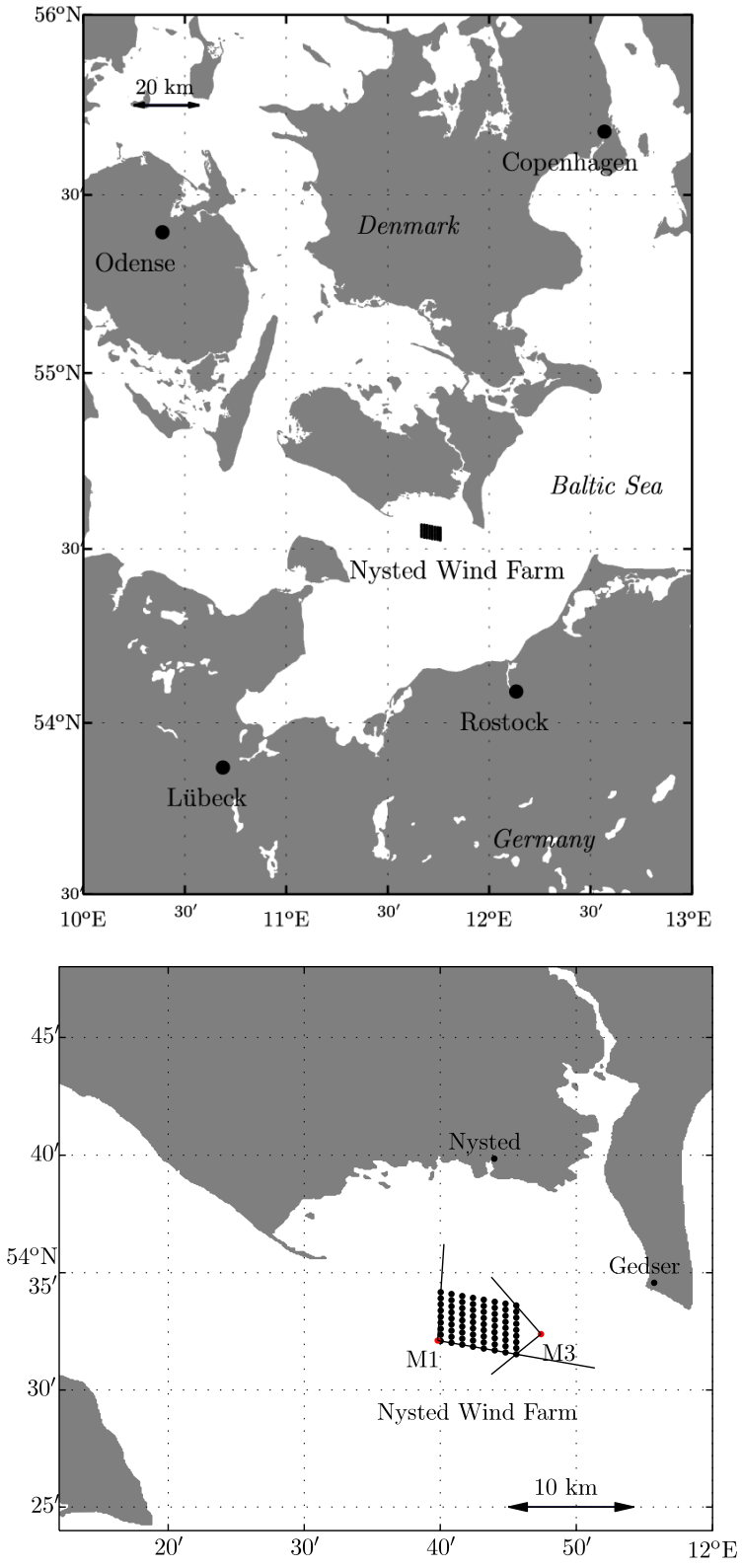


Figure 2.2: The layout and location of the Nysted wind farm with respect to the nearby coastlines. Each black dot within the wind farm area indicates one turbine. The meteorological masts are shown as red dots, labelled M1 and M3. The wind direction sectors where the masts are under the influence of the wind farm wake are shown in the upper plot.

2.3 Time series observations at Horns Rev and Nysted

At Horns Rev M2, the following measurements were available:

- Cup-anemometer wind speed at 62 m, 45 m, 30 m and 15 m
- Wind direction at 60 m, 43 m and 28 m
- Sonic-anemometer wind speed components at 50m (10 month period in 2004 only)
- Air temperature at 55 m and 13 m
- Water temperature at a depth of 4 metres
- Precipitation
- Air pressure at 55 m

At Nysted M1 and M3, the following measurements were available:

- Cup-anemometer wind speed at 69 m, 65 m, 55 m, 40 m, 25 m, 10 m
- Wind direction at 65 m
- Air temperature at 65 m and 10 m
- Water temperature at a depth of 2 metres
- Precipitation
- Air pressure at 10 m

As a result of speculation over an offset in the wind direction measurements from Horns Rev M2, the direction measurements were analysed against a reference data set measured at a height of 10 metres at a coastal point about 20 km from M2, Blåvandshuk. The reference data set was measured by an automatic weather station, and was available as time series observations from 2003–2006, and in raw METAR format from 2000–2003. This meant that 7 years of data was available for comparison. To support the analysis, wind speed measurements from the nacelles of the wind turbines, which were available in 2005 only, were used to estimate the true wind direction, considering that the maximum wake deficit between a pair of turbines should occur when the wind is blowing exactly along the line connecting the two turbines. Using the wake deficit method, it was shown that the direction measurements were reasonable for 2005. The reference data from Blåvandshuk was then used to show that the offset in the data did not change considerably between 2000 and 2006, thus indicating that the errors in the directions throughout the 7 year period were less than about 5 degrees.

2.4 Meteorological conditions over the North Sea

The average meteorological conditions over the North Sea have been addressed in a number of studies. For example, Coelingh et al. (1992, 1996) analysed measurements from offshore sites near the Dutch North Sea coast. They calculated the Obukhov length for their test sites, and found that overall, the area was dominated by unstable atmospheric conditions, with around 70% of days classified as very unstable or unstable in autumn and winter and around 45% in spring and around 55% in summer. The strongest wind speeds were found

when the wind was from the southwest to westerly directions, and the strongest wind speeds were found during the late autumn and winter season of the year. There was little diurnal variation in wind speed, in contrast to later results presented in Coelingh et al. (1998) that showed a distinct diurnal cycle for nearby sites over the land.

Average wind conditions for the two sites are shown in figures 2.3 and 2.3, using data from 2000–2003 for Horns Rev and from 2005–2006 for Nysted. At Horns Rev, the most common wind directions are from the south-southwest to northwesterly directions, and the highest proportion of wind speeds over 15 m s^{-1} are also found in this direction range. At Nysted, most common wind directions are from the south-southwest to westerly directions. The mean wind speed and fitted Weibull distribution for the two sites is also shown on the histograms, where it is seen that the mean wind speed at Horn Rev (9.3 m s^{-1}) is slightly higher than that at Nysted (8.4 m s^{-1}).

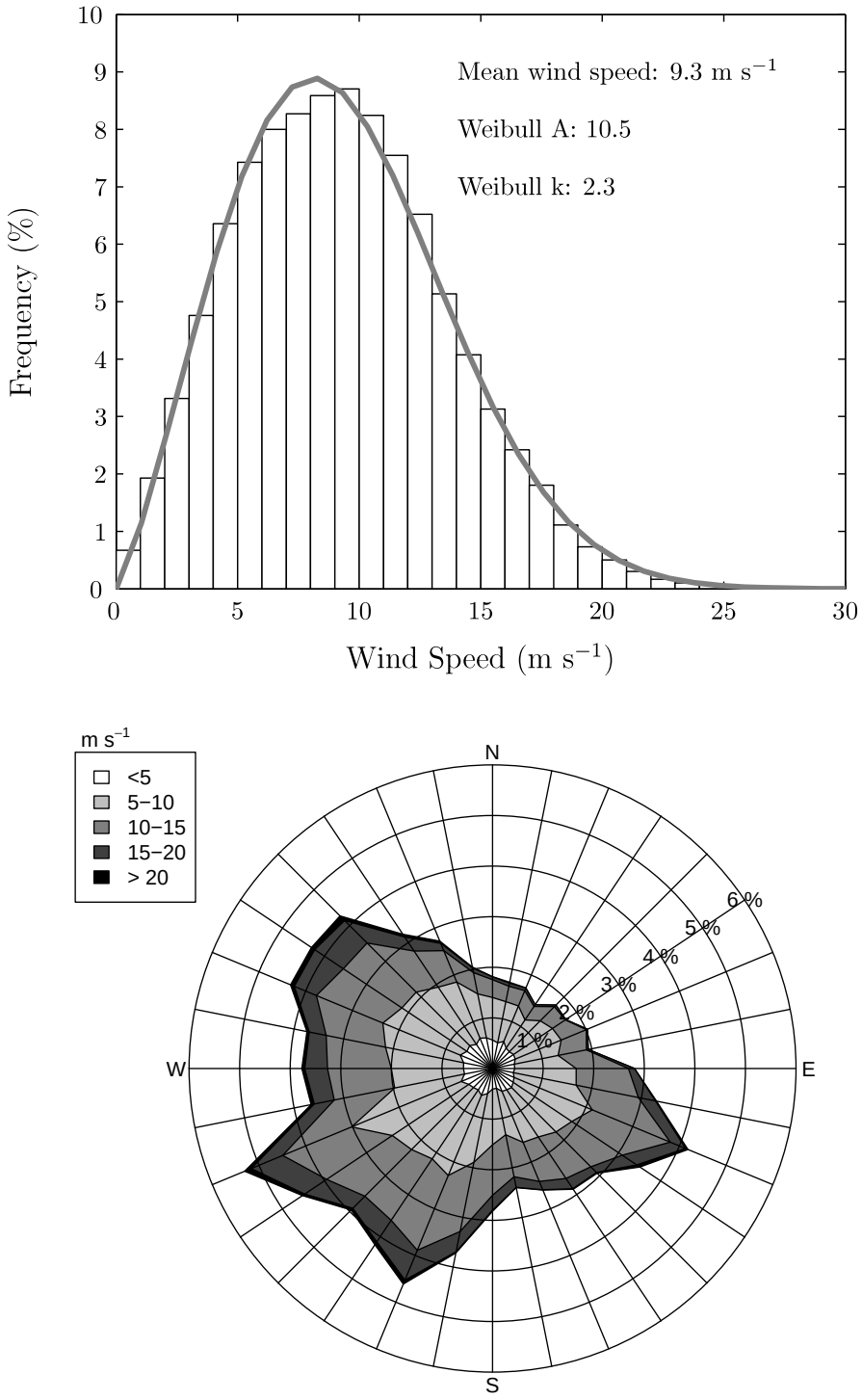


Figure 2.3: Wind rose and histogram for Horns Rev mast 2, for the years 2000–2003. Wind speed measurements are from the top-mounted cup-anemometer of height 62 m. Wind direction measurements are from the wind vane at a height of 43 m. All measurements, including those in the wake of the wind farm, are used.

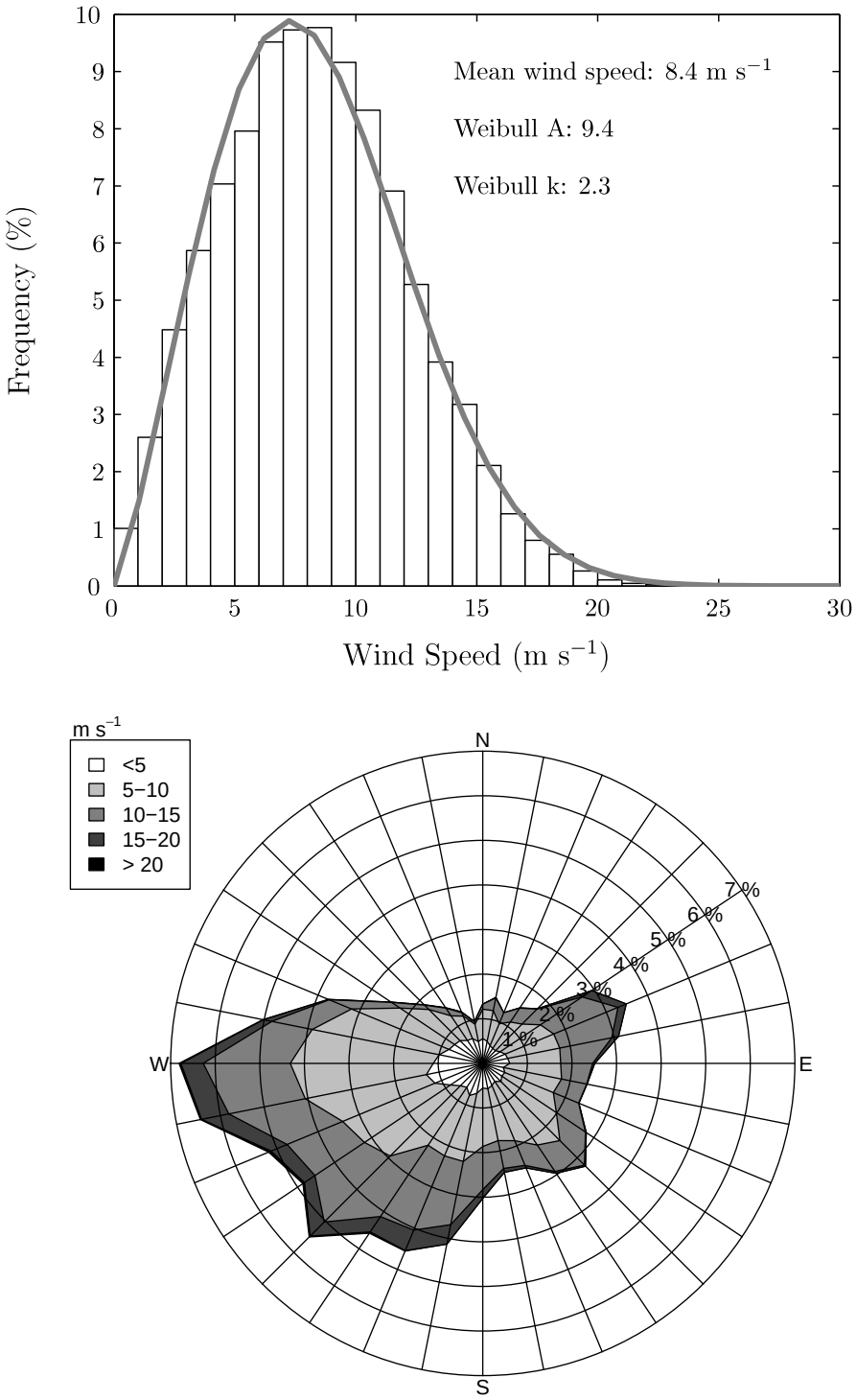


Figure 2.4: Wind rose and histogram for Nysted mast 3, for the years 2005–2006. Wind speed measurements are from the top-mounted cup-anemometer at a height of 69 m. Wind direction measurements are from the wind vane at a height of 65 m. All measurements, including those in the wake of the wind farm, are used.

Chapter 3

Analysis of mesoscale wind fluctuations in wind speed time series

3.1 Introduction

Large fluctuations in the wind speed with a period of around 1 hour can easily be seen in time series observations from meteorological masts near the Horns Rev wind farm. Qualitatively, episodes of severe wind fluctuations can be identified and related to the meteorological situation. An example of an episode of large fluctuations in wind speed that was observed at Horns Rev during the period 22–25 February 2002 is shown in figure 3.1.

To conduct a systematic and statistical study of the occurrence of severe wind fluctuations, however, it is necessary to define a quantitative measure of the instantaneous amplitudes and frequencies of the wind speed. These issues were addressed in Vincent et al. (2010a), where the Hilbert-Huang Transform (HHT) was used to describe the time evolving variability of

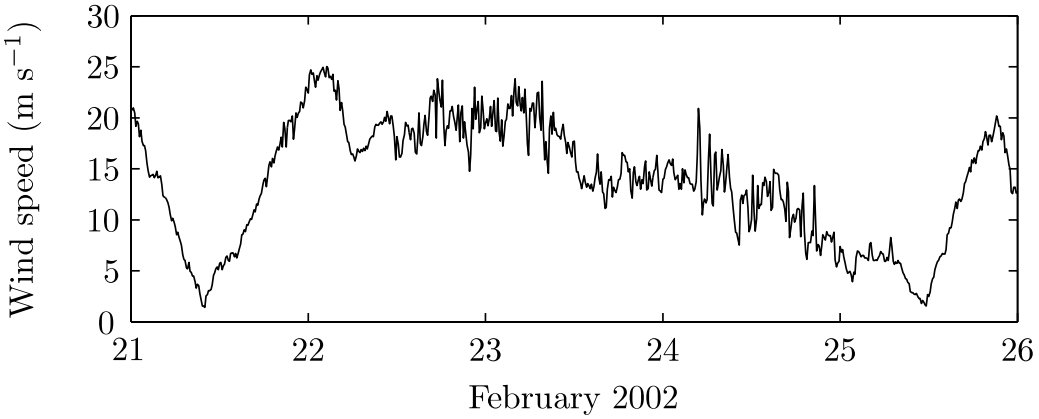


Figure 3.1: Observed wind speed at the 62 m top-mounted cup-anemometer at Horns Rev during the period 22–25 February 2002.

the wind speed observed at the Horns Rev wind farm. Key aspects of the work in Vincent et al. (2010a) are reproduced here, including an overview of the methodology of the HHT and more in-depth arguments relating to the non-stationary of wind speed time series. In this chapter, some additional theoretical and real results relating to the use of the HHT for analysis for wind speed time series are also presented.

3.2 Spectral analysis of non-stationary time series

3.2.1 Stationarity

Vincent et al. (2010a) argued that wind speed time series are statistically non-stationary on a number of time scales. The assumption of non-stationarity indicates that the statistical properties of the wind speed time series evolve with time and justifies the use of non-parametric and adaptive spectral analysis. Most statistical tests for stationarity include some choice of a window over which the stationarity is to be considered, and this limits their applicability for complex time series such as wind speed, where there is no obvious window over which the time series should be stationary. However, wind speed time series must be non-stationary due to the different weather regimes that follow the large scale passage of Rossby waves, the impact of different stability conditions and the onset and decay of various mesoscale phenomena. Even on very long time scales, seasonal, annual and decadal cycles and climate change ensure that there is no time scale on which the wind speed can definitely be considered stationary.

3.2.2 Adaptive spectral analysis

The Fourier spectrum is a commonly used and effective tool for describing the amplitude of fluctuations in a time series as a function of frequency. The Fourier transform finds the best globally fitting set of sine and cosine harmonics to the time series, so a persistent and regular fluctuation on a certain time scale (such as the diurnal cycle or the synoptic cycle) will appear as a peak in the Fourier spectrum.

However, as discussed above, wind speed time series are assumed to be non-stationary, which means that time evolving spectral information will not be well described by a global set of sine and cosine harmonics.

There are a few different methods for adaptive spectral analysis. For example, a moving Fourier transform could be used, where the Fourier transform is applied to a window of length L that slides along the time series. The limitation to this method is that there is an inevitable compromise between the resolution and the length of the window, since a short window will give good resolution but only cover a small range of frequencies, while a long window can cover a wider range of frequencies, but only give coarse resolution.

Another commonly used technique is wavelets, where a wavelet function that can be scaled to represent different frequencies is projected onto the data to find the most important frequencies at each time (Labat, 2005). Wavelets have been used effectively in many different

fields, including to analyse geophysical time series such as turbulence measurements (eg. Barthlott et al. (2007)), and are able to capture the spectral behaviour of non-stationary time series as a function of time. A disadvantage of the method is that an *a priori* decision must be made about the shape of the wavelet function.

Like wavelets, the HHT, which was first introduced by Huang et al. (1998), describes the spectral behaviour of a time series as a function of time. It consists of an empirical filter that decomposes the data into a linear combination of a set of basis functions called intrinsic mode functions (IMFs), followed by extraction of the instantaneous amplitudes and frequencies of each component using the Hilbert transform. The decomposition is called Empirical Mode Decomposition (EMD), and is adaptive and based on the data itself.

3.2.3 The Hilbert Huang Transform

The theory of the normalised HHT, and its implementation, will be reviewed briefly here. Several other detailed descriptions of the method already exist, including (Huang, 2005b) and (Huang and Wu, 2008).

The normalised HHT consists of three steps: EMD of the time series into IMFs, normalisation of the IMFs and extraction of instantaneous amplitudes, and finally extraction of instantaneous frequencies from the normalised IMFs using the Hilbert transform.

The EMD begins by defining two cubic splines. One passes through all the local maxima of the data, and the other passes through all the local minima. The average of the two splines is considered the local mean of the data, and is subtracted from the original time series. The result has zero mean, but does not necessarily form an IMF since it can still contain introduced local extrema. The processes is repeated until convergence is obtained, at which point it satisfies the conditions of being an IMF as defined by Huang et al. (1998). In practice, convergence is defined as in (Huang et al., 2003), as the point where every local maximum–minimum pair is separated by a zero crossing, and when the number of zero crossings stays constant for S iterations. S is termed the stoppage, and is recommended to be a small number between 3 and 8.

When the first IMF, $x_1(t)$, has been calculated using the above procedure, it is subtracted from the original signal:

$$U_1(t) = U(t) - x_1(t). \tag{3.1}$$

$U_1(t)$ is the same as $U(t)$, but has had the highest frequency fluctuations filtered from it. The next IMF is then extracted from $U_1(t)$ using the same process, so that:

$$U_2(t) = U_1(t) - x_2(t). \tag{3.2}$$

When the signal, $U(t)$, has been decomposed into its constituent IMFs, $x_i(t)$, it may be written as:

$$U(t) = \sum_{i=1}^{N} x_i(t) + \varepsilon(t) \tag{3.3}$$

where N is the number of IMFs that have been extracted from the signal and $\varepsilon(t)$ is the low frequency trend, containing zero or one extrema.

Correct treatment of end effects in the decomposition is an ongoing question, since it involves making a reasonable prediction of the spline from the last extremum to the end of the time series (Huang and Wu, 2008). In this work, however, the time series are very long and we are interested in average variability conditions, so the end effects do not present any significant problem. For other applications, the end effects problem might become critical. For example, if the Hilbert-Huang transform were to be used as a forecasting tool, as has been proposed by Kim et al. (2008), then any use of autoregressive components in a statistical model would rely on a reasonably accurate estimate of the variability or the individual IMFs very close to the end of the time series. This issue is discussed further in chapter 7, where several statistical models of wind variability based on input from the large scale flow and prior values of the wind variability itself are implemented.

In a recent modification to the HHT, Huang (2005a) proposes normalizing the IMFs so that their amplitude is always unity and they contain only frequency modulations. The normalised HHT avoids the potential problem that if the spectrum of oscillations within the IMF and that of their low frequency envelope overlap, then the instantaneous frequency will not be meaningful according to the Bedrosian theorem (Huang and Wu, 2008). Each IMF is normalised by dividing it by the cubic spline that passes through the absolute value of all its extrema. The normalization should be repeated iteratively until all extrema are equal to unity, because the spline can occasionally pass through values less than the IMF. Each IMF, $x_i(t)$ may then be divided into a frequency modulation part, $F_i(t)$ and an amplitude modulation part, $A_i(t)$, where $F_i(t)$ is the normalised IMF with amplitude of unity, and

$$A_i(t) = \frac{x_i(t)}{F_i(t)} \tag{3.4}$$

The normalised IMFs satisfy the special property that they each contain only one frequency at each time, so that instantaneous frequencies can be calculated using the Hilbert transform. The meaning of instantaneous frequency is not immediately obvious. Conceptually, it could be regarded as the inverse of the time taken for one complete oscillation, but as argued by Huang et al. (2009), there is no reason to expect that the frequency should remain constant throughout an entire oscillation. The appropriate definition of the local frequency is as the derivative of the phase angle, where the phase angle can be found by adding some properly chosen imaginary part to the time series.

$$f(t) = x_i(t) + ig(t) \tag{3.5}$$

such that

$$x_i(t) = \mathrm{Re}\left[f(t)\right]. \tag{3.6}$$

Clearly, any function $g(t)$ satisfies Eqs. (3.5) and (3.6). In general, however, the complex signal $f(t)$ will contain both positive and negative frequencies, and the frequencies will not have physical meaning. As discussed by Cohen (1989), a breakthrough in the idea of defining an instantaneous frequency was made by Gabor (1946) when a special function $g(t)$ was used which filtered out the negative frequency components and doubled the positive frequency components. This function was equivalent to the Hilbert transform, and is given as

$$H_i(t) = x_i(t) + i\frac{1}{\pi}\mathrm{PV}\int_{-\infty}^{+\infty}\frac{x_i(\tau)}{(t-\tau)}d\tau = x_i(t) + ih_i(t) \tag{3.7}$$

where PV refers to the principal value of the integral, which must be considered because the integrand is not defined at $t = \tau$.

Mathematically, the Hilbert transform of the original data could be calculated, but the result would be erroneous to interpret because the instantaneous frequency would be a result of all the overlapping frequencies present in the original signal (Boashash, 1992). The construction of the normalised IMFs ensures that each component satisfies the condition of "monocomponent signals" as suggested by Cohen (1995) and Boashash (1992), which means that the spectrum of the time varying amplitude does not overlap with the spectrum of the time varying phase. In a practical sense, each normalised IMF should contain only one frequency at any time.

By writing the real and complex parts of the Hilbert transformed signal in polar coordinates, it is seen to be function of instantaneous phase and amplitude,

$$H_i(t) = a_i(t)e^{i\theta_i(t)}. \tag{3.8}$$

where in the case of the normalised IMFs, the amplitude will be equal to unity. Defining the instantaneous frequency as $\omega_i(t) = \frac{d\theta_i(t)}{dt}$, the Hilbert transform may then be written as

$$H_i(t) = a_i(t)e^{i\int \omega_i(t)dt}. \tag{3.9}$$

Since the instantaneous phase may be expressed as $\theta(t) = \mathrm{Im}\left[\ln(H(t))\right]$, the instantaneous frequency, defined as the time derivative of the instantaneous phase, can be written as

$$\omega_i(t) = \frac{d\theta_i}{dt} = \mathrm{Im}\left[\frac{1}{H_i(t)}\frac{dH_i(t)}{dt}\right]. \tag{3.10}$$

The data is discrete, so the instantaneous frequency is then approximated as

$$\omega_i(t) \approx \mathrm{Im}\left[\frac{1}{H_i(t)}\frac{\Delta H_i(t)}{\Delta t}\right]. \tag{3.11}$$

After the Hilbert transform of the N IMFs has been established, the original signal may be reconstructed as

$$U(t) = \mathrm{Re}\left[\sum_{i=1}^{N} a_i(t)e^{i\int \omega_i(t)dt}\right] + \varepsilon(t). \tag{3.12}$$

where $\varepsilon(t)$ is the low frequency trend in the data after the first N IMFs have been filtered out.

Eq. (3.12) shows how the signal may be decomposed into a series of IMFs, each with time varying amplitude and phase. An example of the first five IMFs of such a decomposition for a 20 day sample of wind speed observations is given in figure 3.2. As discussed by Sweeney-Reed and Nasuto (2007) and Wu and Huang (2009), direct physical interpretation of a single IMF is not necessarily possible, due to the problem of mode mixing, where an IMF can contain parts of fluctuations belonging to different scales. For example, part of the diurnal cycle may appear at some times in the same IMF as parts of the synoptic cycle at other times. This problem can be observed in some of the IMFs shown in figure 3.2 - for example, the fluctuations in IMF 4 on day 14 are slower than the fluctuations in IMF 5 on days 16–17. To create IMFs which have a more direct physical meaning (as well as greater uniqueness), Wu and Huang (2009) developed the Ensemble Empirical Mode Decomposition (EEMD) methodology, where white noise is added to the time series to create an ensemble

of decompositions. The white noise ensures that all frequencies are present at all times in the time series, and removes the problem of mode-mixing. By averaging the ensemble of decompositions, a true decomposition is found which has greater physical meaning, but for which the components may not exactly satisfy the conditions of being IMFs. The EEMD technique has not been pursued here, but it would nonetheless be an important extension to the current work, particularly if unique physical processes (such as oscillations associated with roll vortices) were to be identified in a single IMF.

An alternative to using the Hilbert transform for calculation of instantaneous frequency is to use the direct quadrature, as discussed in (Huang et al., 2009). This method takes advantage of the fact that the phase can be calculated directly as the arc-tangent of the frequency modulation part of the signal divided by its quadrature. Huang et al. (2009) argues that it is a more local calculation of the instantaneous frequency, and that it avoids any remaining violations of the Bedrosian theorem.

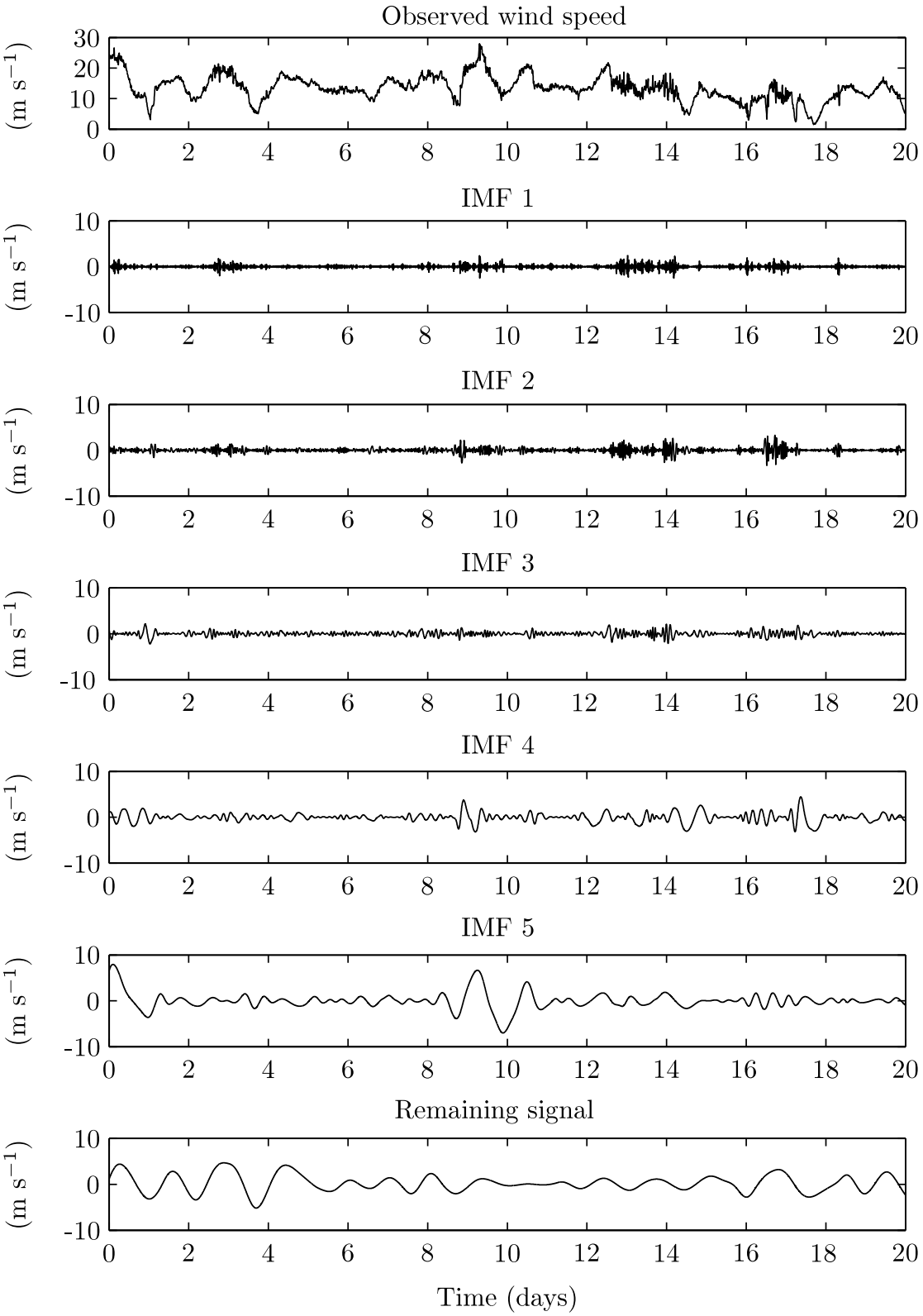


Figure 3.2: The first five IMFs from the Empirical Mode Decomposition of a wind speed time series. The top panel shows the corresponding wind speed observations, and the bottom panel shows the remaining signal after the first five IMFs have been subtracted from the time series.

One way of combining all of the frequency and amplitude information in the set of IMFs is to bin the instantaneous amplitude information in each IMF into regularly spaced frequency bins. This can be achieved by considering each frequency bin $(\omega, \omega + \Delta\omega]$ and equating all amplitude contributions $a_i(t)$ from the N IMFs which correspond to $\omega_i(t) \in (\omega, \omega + \Delta\omega]$. The sum of the total amplitude contributions in each frequency bin at each time constitute the Hilbert Spectrum, $H(\omega, t)$. Other methods for handling this step include using non-parametric regression techniques to fit a curve through the amplitudes, which also has the advantage of smoothing the spectrum (Duffy, 2004). However, since in this study the Hilbert spectrum is finally averaged or converted to a scalar time series (to which is it easy to apply smoothing), no smoothing was applied to the raw Hilbert spectrum. The Hilbert spectrum has the same units as the original data set, as it represents amplitudes of the fluctuations.

3.3 Examples of the Hilbert Huang transform as a tool for analysing wind speed time series

Important requirements of the analysis methodology are that it responds quickly to sudden changes in the statistical properties of the time series and that it differentiates effectively between different frequencies. The extent to which the Hilbert spectrum satisfies this requirement is demonstrated using both a theoretically constructed and a real time series of wind speed. The theoretically generated time series was designed to include sudden and gradual changes in both frequency and amplitude, and was constructed as a combination of sine curves, with random noise of mean 0.5 m s^{-1} and standard deviation 0.3 m s^{-1} then added to the whole time series. The time series was 2880 time steps long, which is equivalent to 20 days of observations with a 10 minute sampling frequency.

The time series and its Hilbert spectrum are given in figure 3.3, where the time dimension of the time series is considered in units of days. In this figure, the Hilbert spectrum was multiplied by the frequency to emphasise the high frequency components of the fluctuations, so the units are m s^{-2}. It is seen that the fluctuations on day 2, which have a period of 5 hours, are represented on the Hilbert spectrum with a dark contour of frequency 0.2 cycles hr^{-1}, while the subsequent fluctuations between days 3 and 5, which have a period of 8 hours, are represented by a dark contour of 0.125 cycles hr^{-1}. Between days 5 and 10, the fluctuations are not shown on the Hilbert spectrum because they are slower than the lower frequency axis limit of 10 cycles hr^{-1}. On day 10, the gradually decreasing period is seen on the Hilbert spectrum as a line with positive gradient. Finally, the episode of high frequency fluctuations of period 1.5 hours are seen as an area of dark contours between about 0.45 and 0.6 cycles hr^{-1}. These frequencies spread over are larger range and are slightly incorrect, because the fluctuations are close to the sampling frequency of 10 minutes.

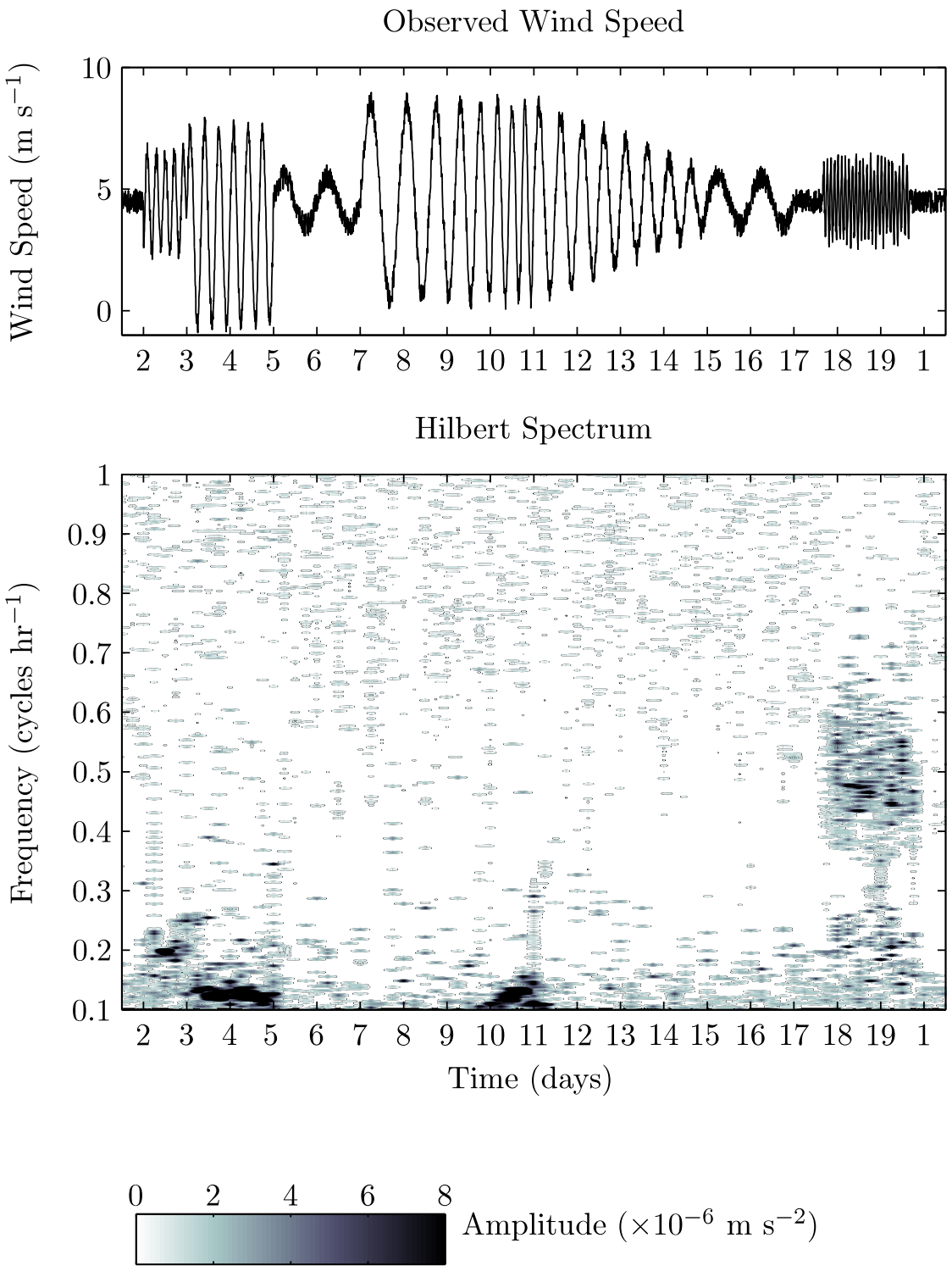


Figure 3.3: Theoretical time series (top panel) and its corresponding Hilbert Spectrum (bottom panel)

A real time series of observed wind speed at Horns Rev, with the same length and sampling period as the theoretical time series, is shown in figure 3.4, together with its Hilbert spectrum.

The episodes of enhanced mesoscale variability on days 3–4, 13–16 and 17–18 are seen as dark contours on the Hilbert spectrum.

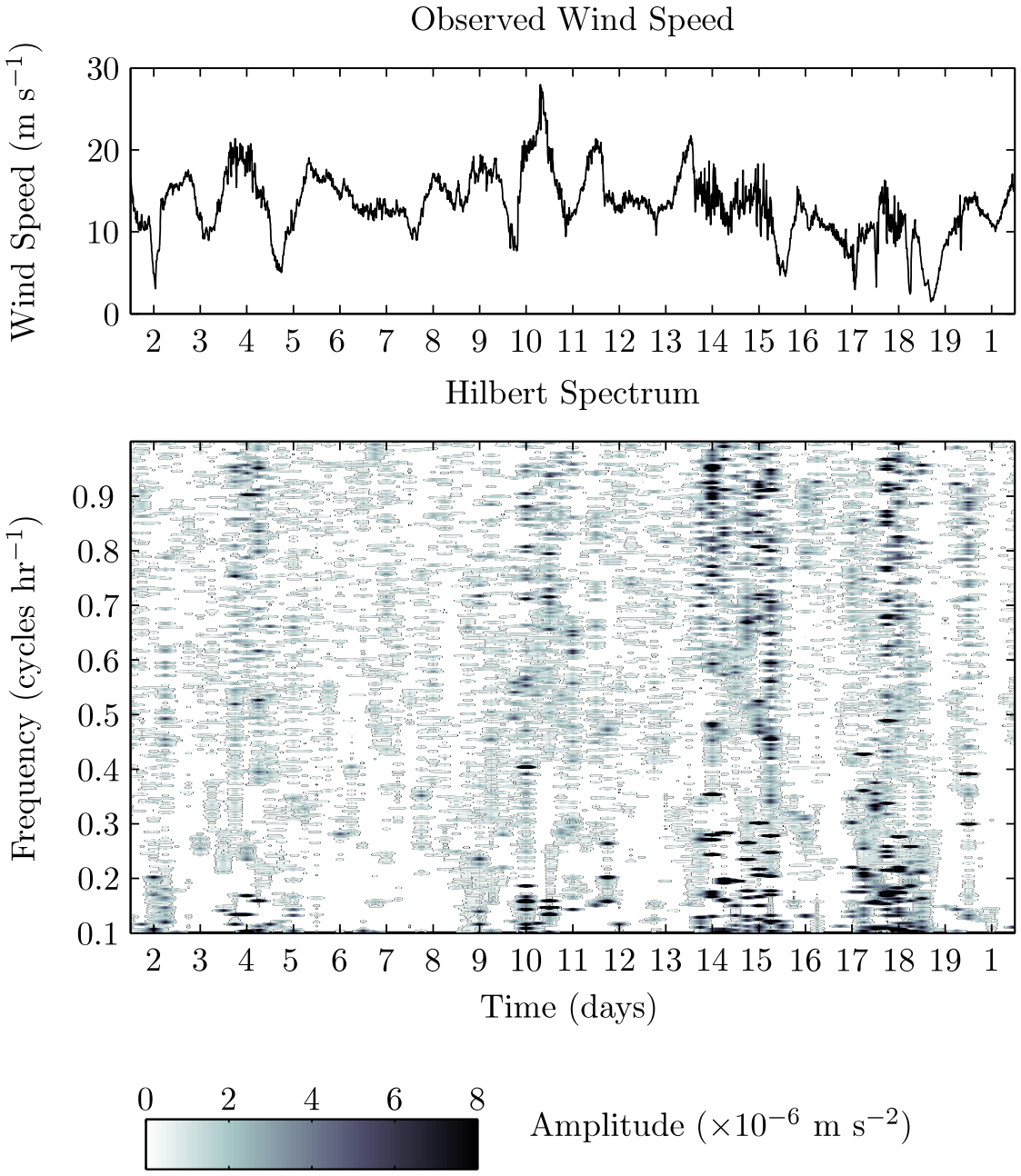


Figure 3.4: Observed time series (top panel) and its corresponding Hilbert Spectrum (bottom panel)

While the Hilbert spectrum is a useful way of visually assessing the spectral information in a time series, for practical purposes a scalar metric of the degree of variability on a certain time scale is a useful tool for further analysis. Such a time series can be calculated by integrating the Hilbert spectrum between two time scales, as given in equation 3.13.

$$V(t) = \sum_{\omega=\omega_1}^{\omega_2} H(\omega, t) \tag{3.13}$$

where $V(t)$ represents the sum of all amplitudes within the frequency range $[\omega_1, \omega_2]$.

Due to the superposition of fluctuations of different phase and frequency, the signal will not always be oscillating with this total amplitude. The amplitude can be interpreted as an upper bound on the total amplitude of fluctuations if all components were in phase, and combined according to linear superposition, although the assumption of linear superposition probably does not always hold. The value of a scalar metric of variability is that it can be related directly to the prevailing meteorological situation, predicted using univariate time series modelling or using various explanatory variables, or used as an input to down-stream power system or wind farm models. The predictability of the variability time series based on the large scale meteorological conditions will be further addressed in chapter 7.

The variability metric for the time series given in figure 3.3 is shown in figure 3.5 for five different time scales. For the fastest periods of 1–3 hours, the only peak in the variability time series is during the simulated episode of variability on days 17–18. For the part of the time series where the period is gradually increasing (between days 6 and 10), the variability signal is first identified in the 18–30 hour variability time series, and subsequently in the 10–18 hour and 6–10 hour variability time series. For the part of the time series with a constant period of 12 hours but decreasing amplitude (days 10–14), the signal is identified in the 10–18 hour variability time series, with the negative slope indicating the decreasing amplitude. The graphs indicated that this analysis methodology is ideal for the analysis of mesoscale wind fluctuations, where only the fluctuations of a certain time scale are of interest.

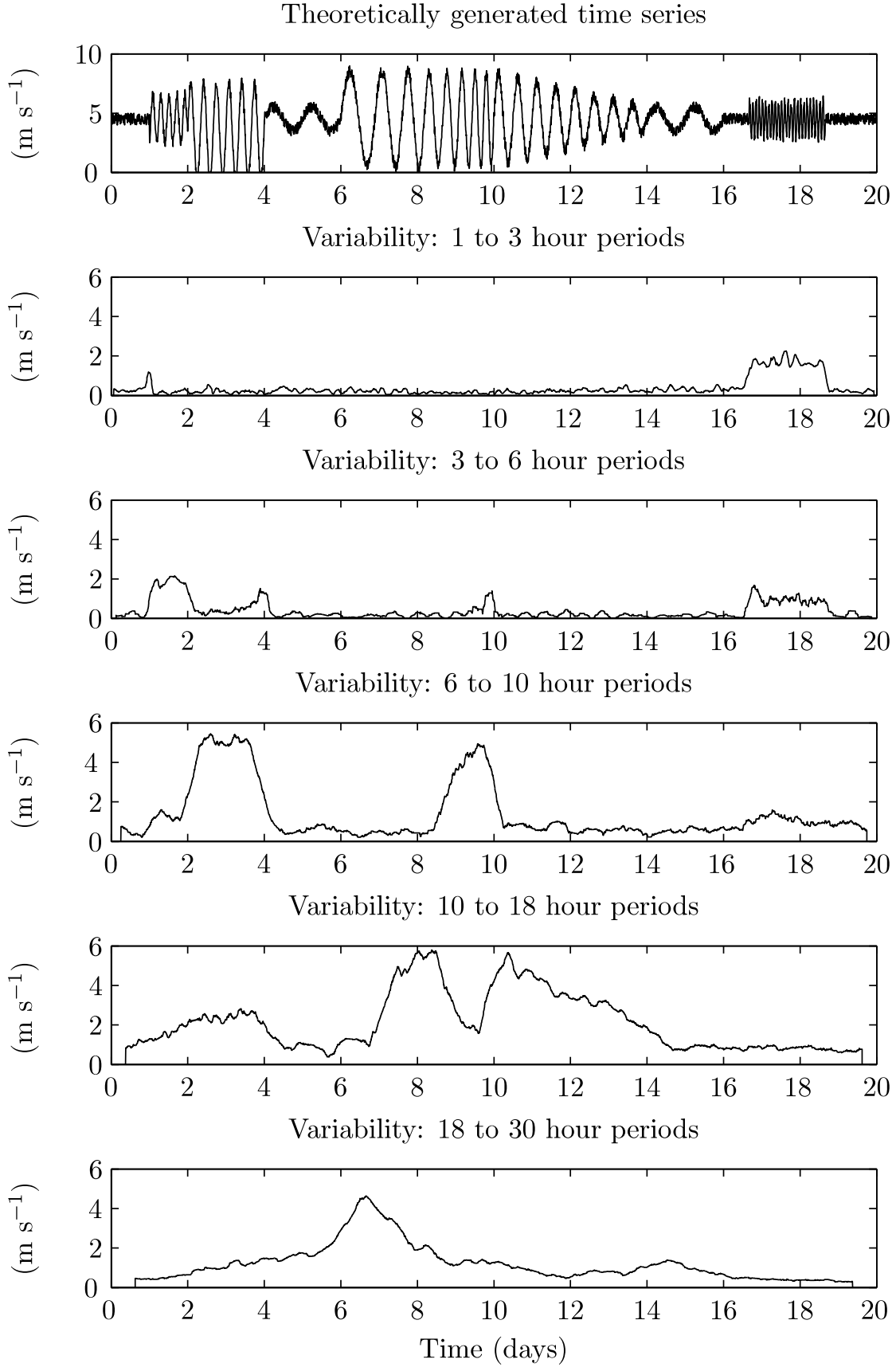


Figure 3.5: Theoretical time series (top panel), and variability time series for five different time scales: 1–3 hour periods, 3–6 hour periods, 6–10 hour periods, 10–18 hour periods, and 18–30 hour periods.

Another important application of the Hilbert spectrum is to bin and average the data along the time axis to create conditional spectra. This is a particularly interesting application of the method, since it permits calculation of average spectra for points which do not lie consecutively in the time series, such as all the points where the wind direction falls within a certain range. In this way, the horizontal time axis can be remapped to some other parameter. Further, the fast adaptivity of the method means that the spectral information for an isolated point in the time series is likely to be reasonably free from the influence of the spectral information from nearby parts of the time series. This method was used extensively in Vincent et al. (2010b), and will be applied to the analysis of time series from the Horns Rev and Nysted wind farms in chapter 4.

3.4 Discussion and Conclusions

Results presented in this chapter show that the HHT is a useful tool for the analysis of wind speed time series. It can respond to the rapidly changing statistical properties of the wind speed, which is an important property for analysing the dependence of wind variability on changing weather conditions.

Three different ways of presenting the information in the Hilbert spectrum were discussed. Firstly, the contour plot of the Hilbert spectrum as a function of frequency and time is an interesting way to display the frequency information contained in a relatively short time series. Secondly, the time series of total variability on certain time scales, which is calculated by integrating the Hilbert spectrum over a range of frequencies, provides a useful scalar index of wind variability. Thirdly, conditional averaging along the time axis of the Hilbert spectrum re-maps the horizontal axis from time to some other parameter of interest.

There are several shortcomings of the HHT. One is that no robust way of dealing with end effects exists. This is only a minor disadvantage in the work presented here, since the time series are very long and any end effects will have an insignificant effect on the whole spectrum. A second problem is that of 'mode mixing'. Since the IMFs extract the frequencies that exist in the time series at each given time, if a particular frequency is not represented at certain parts of the time series, different frequency ranges can appear in the same IMF, and a physical and persistent frequency (such as the diurnal cycle) could be split between more than one IMF. This problem disappears when the Hilbert spectrum is calculated, but it limits the direct physical interpretation of the IMFs themselves. Direct physical interpretation of the IMFs is an interesting prospect for the case-study analysis of particular weather events. The problem of mode mixing is discussed in Wu and Huang (2009), who address the problem by adding N different realisations of white noise to the time series and apply EMD to each of them, a process they call the 'Ensemble Empirical Mode Decomposition' (EEMD). The white noise ensures that all frequencies are represented in the time series at all times, and avoids the problem of mode mixing. The average of the N different sets of IMFs is then considered as the 'true' decomposition. A final shortcoming is that there is no mathematical formulation of the EMD process, which makes it difficult to prove certain aspects of its performance.

Another intriguing application of the HHT that is not explored here is as a forecasting tool. For example, one could use time series of variability on different non-overlapping temporal scales or the IMFs themselves to develop wind forecasting methodologies based

on, for example, vector auto-regressive models as proposed in Kim et al. (2008). This has the interesting advantage of forecasting different modes of variability separately, and then constructing the overall time series out of the different forecast modes.

The HHT is used extensively in Chapter 4 to define time series of mesoscale variability for long time series observations at the Horns Rev and Nysted wind farms, and to calculate spectra conditional on the prevailing meteorological conditions. In chapter 5, it is used to identify severe variability days and relate them to the large scale meteorological conditions and the satellite patterns over the North Sea, and in chapter 8 to identify interesting cases for modelling using the mesoscale model WRF. In chapter 7, the predictability of time series of mesoscale variability using statistical models with the large scale meteorological conditions as input will be explored.

Chapter 4

Mesoscale wind fluctuations at Horns Rev and Nysted

4.1 Introduction

In chapter 3, the Hilbert-Huang Transform (HHT) was introduced as a useful tool for analysing wind speed time series. In particular, it was argued that the HHT could capture the non-stationary aspect of the wind, so that a time-evolving description of the wind speed spectrum could be calculated.

In this chapter, the HHT is applied to time series from meteorological masts near the Horns Rev and Nysted wind farms. Time series of variability for time scales of 1–3 hours are calculated, and conditional spectra based on all available secondary information are presented.

4.2 Time-series of variability

As discussed in chapter 3, a time series of total variability over a range of frequencies can be constructed by integrating the Hilbert spectrum between two frequencies. Time series of variability for fluctuations with periods between 1 and 3 hours were calculated for the 4 years of Horns Rev data and the 2 years of Nysted data that were described in chapter 2. The Horns Rev time series had only a small number of gaps, but the Nysted data had several large gaps that could not be filled in any reasonable way. Therefore, the Nysted time series were analysed in several pieces to avoid the largest gaps.

4.2.1 Horns Rev

The time series of wind speed from the 62 m top-mounted cup-anemometer at Horns Rev M2 for the four year period 2000–2003 are shown in figure 4.1, and the corresponding time series of variability on time scales of 1–3 hours are shown in figure 4.2. The variability time series were smoothed using centred moving average smoothing, with a window length of 3 hours. To show the general trend in the variability time series, they were also smoothed using a window length of 2 days, as indicated by the red curves in figure 4.2.

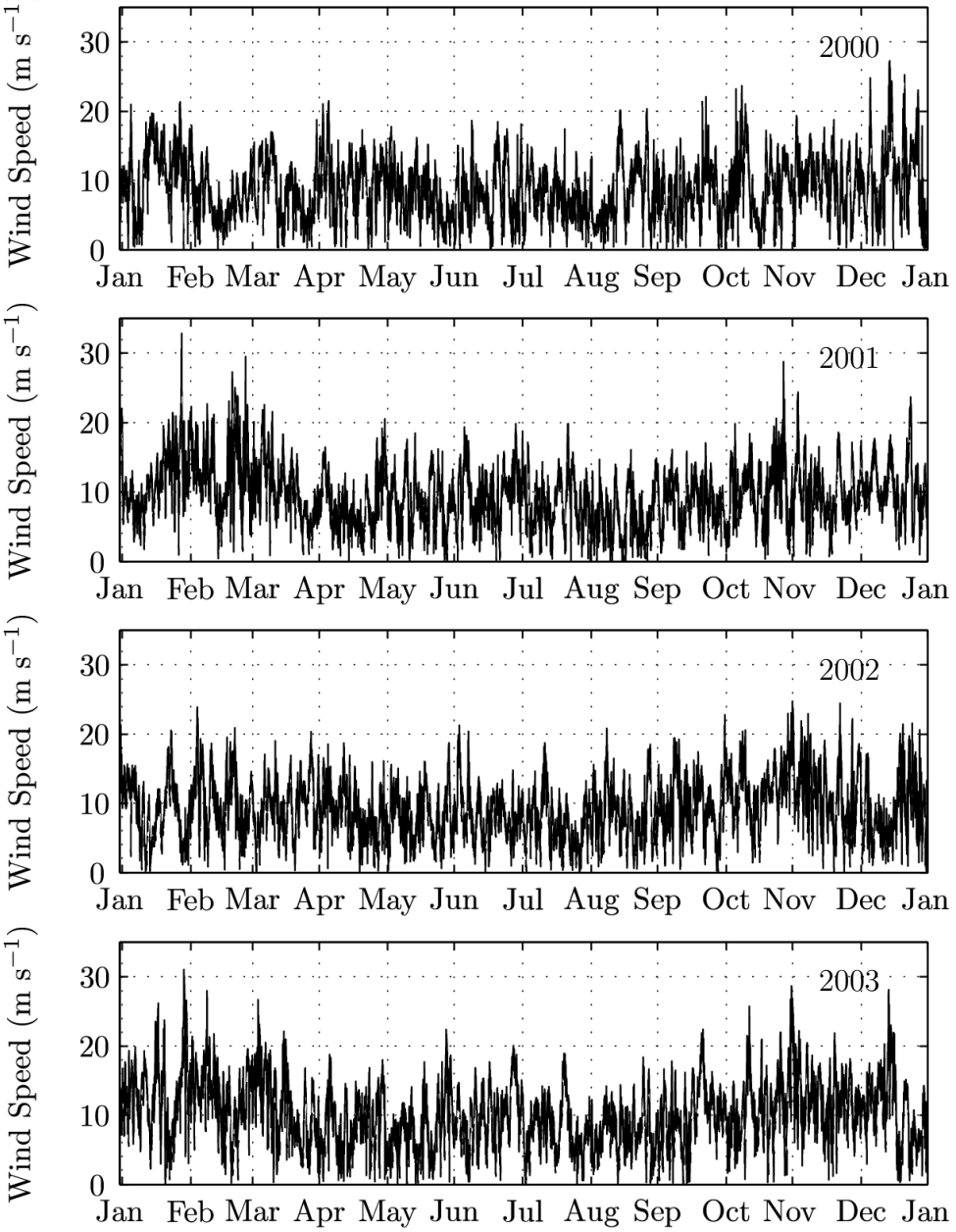


Figure 4.1: Observed wind speed at the 62 m top-mounted cup-anemometer at Horns Rev M2. Top to bottom, 2000–2003.

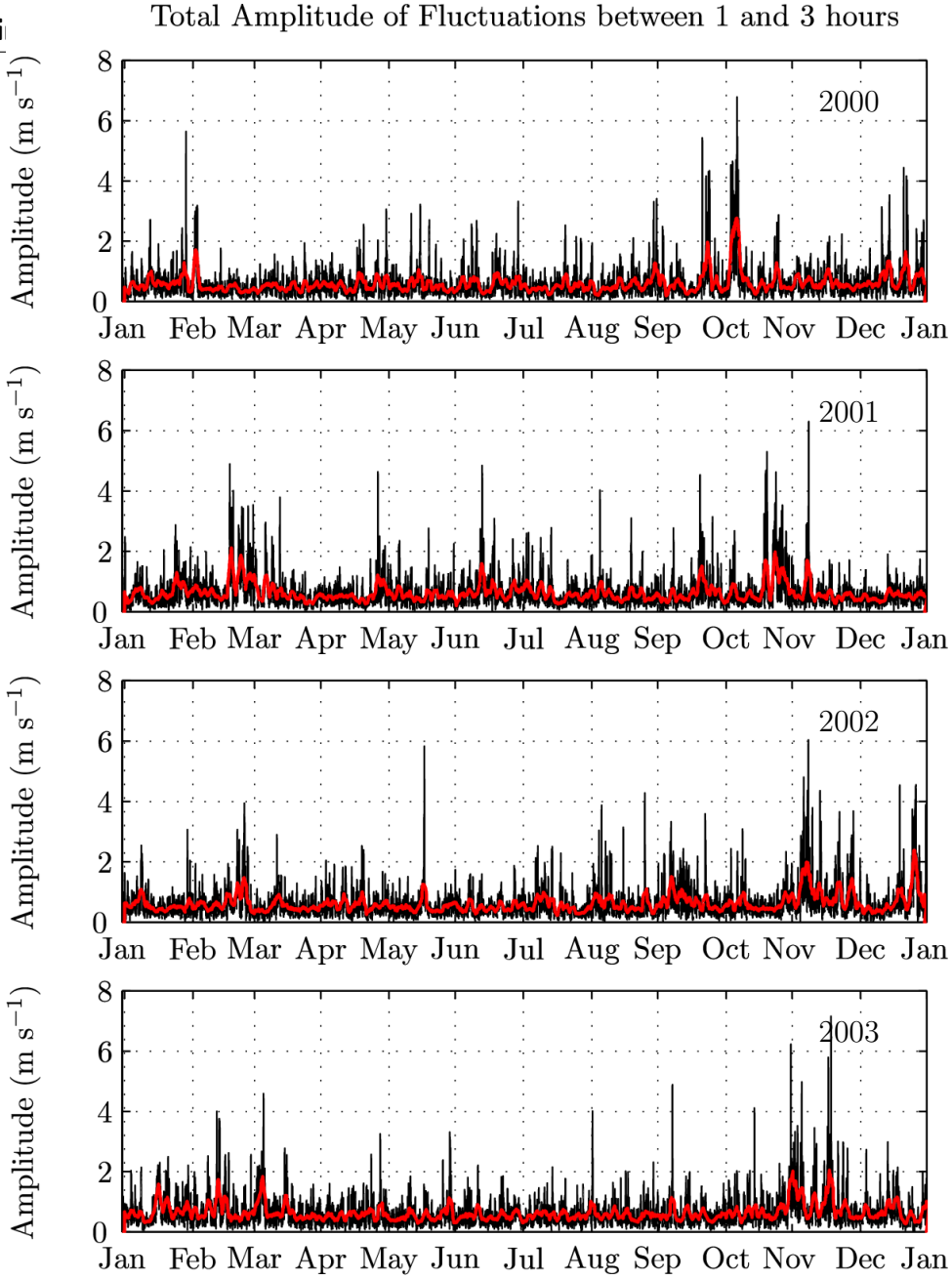


Figure 4.2: Total amplitude of variability for time scales of 1–3 hours, for the observed wind speed at the 62 m top-mounted cup-anemometer at Horns Rev M2. Top to bottom, 2000–2003. Moving average smoothing with a window length of 3 hours was applied to the black curves, and with a window length of 2 days to the red curves.

The time series of variability for Horns Rev indicate that there were periods of enhanced mesoscale wind variability during the months of October to March in each of the four years, with generally lower variability in the months April to September. This trend will be explored further in relation to average Hilbert spectrum.

The time series themselves do not show significant other information about the variability, but they will be used as a building block for further analysis. For example, they will be used to define severe variability days for comparison with synoptic patterns, satellite pictures and wind classes in chapter 5 and for identifying interesting cases to model in chapter 8. The possibility of directly predicting the time series of variability based on the large scale flow

conditions is explored in chapter 7.

4.2.2 Nysted

The time series of wind speed at the 69 m top-mounted cup-anemometer at Nysted M1 for the two year period 2005–2006 are shown in figure 4.3, and the corresponding time series of variability on time scales of 1–3 hours are shown in figure 4.4. The variability time series were smoothed using centred moving average smoothing, with a window length of 3 hours. To show the general trend in the variability time series, they were also smoothed using a window length of 2 days, as indicated by the red curves in figure 4.4.

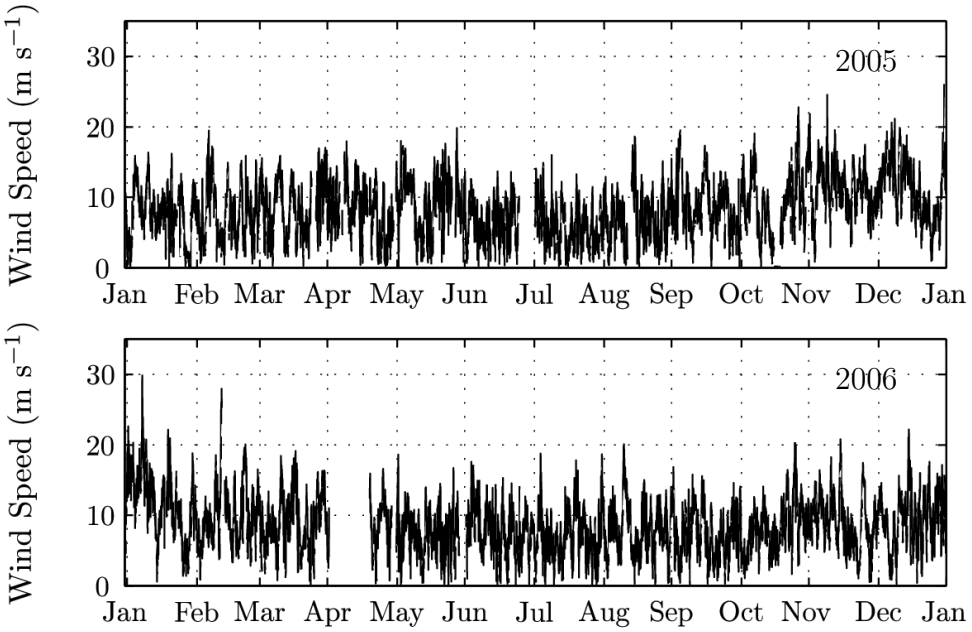


Figure 4.3: Observed wind speed at the 62 m top-mounted cup-anemometer at Nysted M3. Top to bottom, 2005–2006

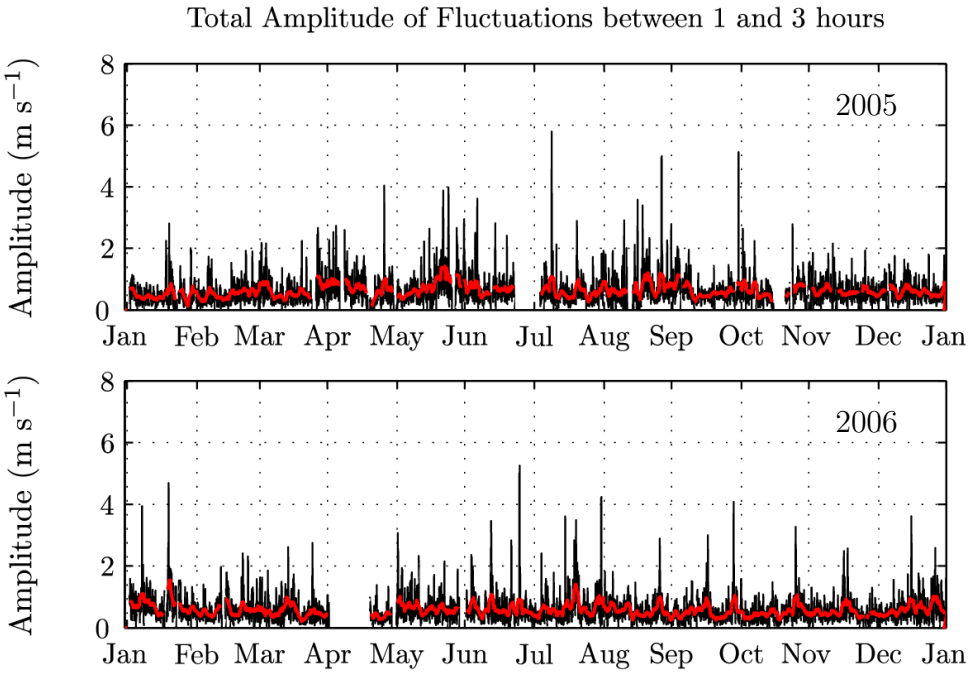


Figure 4.4: Total amplitude of variability for time scales of 1–3 hours, for the observed wind speed at the 62 m top-mounted cup-anemometer at Nysted M3. Top to bottom, 2005–2006. Moving average smoothing with a window length of 3 hours was applied to the black curves, and with a window length of 2 days to the red curves.

The results for Nysted, which are less representative than those for Horns Rev due to the shorter data availability and the large gaps in the time series, do not appear to indicate any obvious annual pattern. This will be demonstrated further through the use of conditional spectra.

4.3 Conditional Spectra for time scales of 1–10 hours

In chapter 3, the possibility of remapping the time axis of the Hilbert spectrum to some other parameter by binning and averaging along the horizontal axis was discussed. This technique is used here to create conditional spectra that show the amplitude of the wind variability as a function of frequency and season, wind direction, wind speed, air-sea temperature difference and pressure tendency, respectively. Similar analyses for Horns Rev and Nysted were conducted, and the equivalent results from the two sites are shown side by side. This emphasises the fact that some results are entirely governed by site-specific effects, while other results may have a more general validity.

4.3.1 Wind variability as a function of time of year

The Hilbert spectra for Horns Rev and Nysted were conditionally averaged according to time of year, using a bin width of 13 days (1872 10 minute observations). For Horns Rev,

the data covered the 4 year period 2000–2003, while the Nysted analysis covered the 2 year period 2005–2006. The amplitude of the Hilbert spectrum was multiplied by the frequency to emphasise the high frequency part of the spectrum, and scaled by the width of the frequency bin used in creating the Hilbert spectrum. Therefore, the units are in m s^{-1}, but should not directly be interpreted as an amplitude of wind speed and have therefore been labelled as 'Variability Intensity'. The conditional spectra for both sites are shown in figure 4.5.

The conditional spectrum for Horns Rev shows increased variability, on average, between the months of October and March, in comparison with the warmer months of the year between April and September. In contrast, the conditional spectrum for Nysted shows generally higher average variability between about April and September, but there is little seasonal pattern for periods shorter than around 2 hours.

The time series of 4 and 2 years for Horns Rev and Nysted, respectively, are not long enough to give a representative annual cycle. This is evident from the time series 1–3 variability graphs at Horns Rev in figure 4.2, where several episodes of strong variability are seen each year. This is reflected in the vertical streaks in figure 4.5, where the average is influenced by a few strong events. Nonetheless, outbreaks of enhanced mesoscale variability occur in each of the four analysis years, and it seems unlikely that this trend is coincidental.

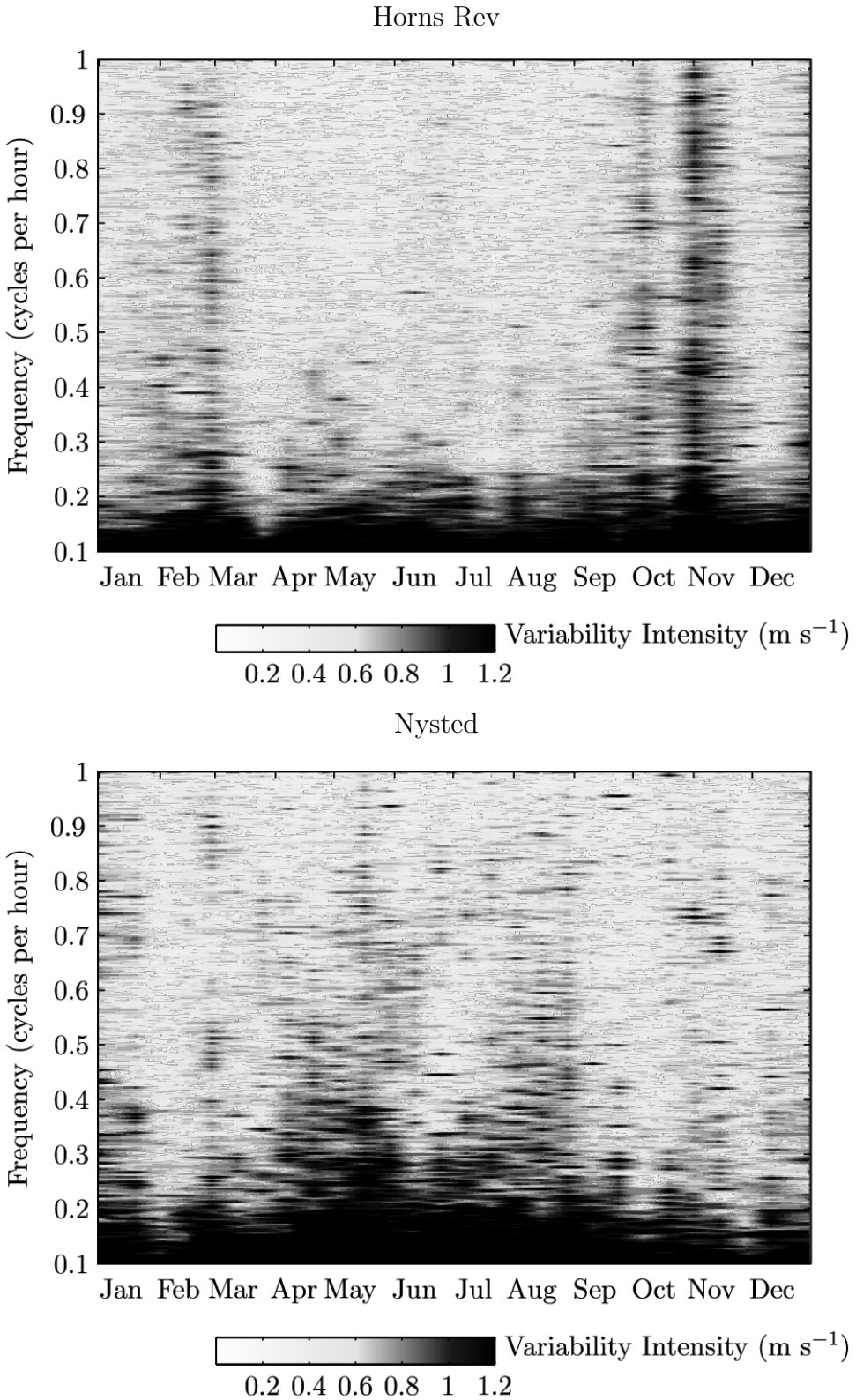


Figure 4.5: The Hilbert spectrum for the observed wind speed at a height of 62 m at Horns Rev M2 (top) and the observed wind speed at a height of 69 m at Nysted M3 (bottom), conditionally averaged according to the time of year

4.3.2 Wind variability as a function of wind direction

Conditionally averaged spectra for Horns Rev and Nysted according to wind direction are shown in figure 4.6. In both cases, a direction bin width of 20 degrees was used. The analysis is for the same 4 year and 2 year time series for Horns Rev and Nysted respectively as for figure 4.5. The analysis was also separated into the 4 seasons of the year (where the seasons are defined as 1 March–31 May for spring, 1 June–31 August for summer and so on) since both synoptic patterns and the interaction of the boundary layer with the land and sea surface change in the course of the year.

The dotted lines indicate the part of the spectrum that could be affected by the wind farm wake in both cases. No obvious signature of the wake is seen at either site. For the Horns Rev case, the flow from the sea and the land is divided between wind directions with a westerly component and wind directions with an easterly component. This is also indicated on the spectrum. For Nysted, the difference between flow from the sea and flow from the land is more complicated (as already discussed in relation to figure 2.2 in chapter 2).

The dependence of wind variability on wind direction is a site specific relationship, since both the surface types and the local wind direction itself are modulated by the configuration of land and sea around the site. One aspect of the dependence that may not be entirely site specific for the northern hemisphere mid-latitudes is that the large scale weather systems tend to approach from the west-northwesterly directions. This may be reflected in the enhanced variability in autumn and winter in the west to north-easterly directions at both sites.

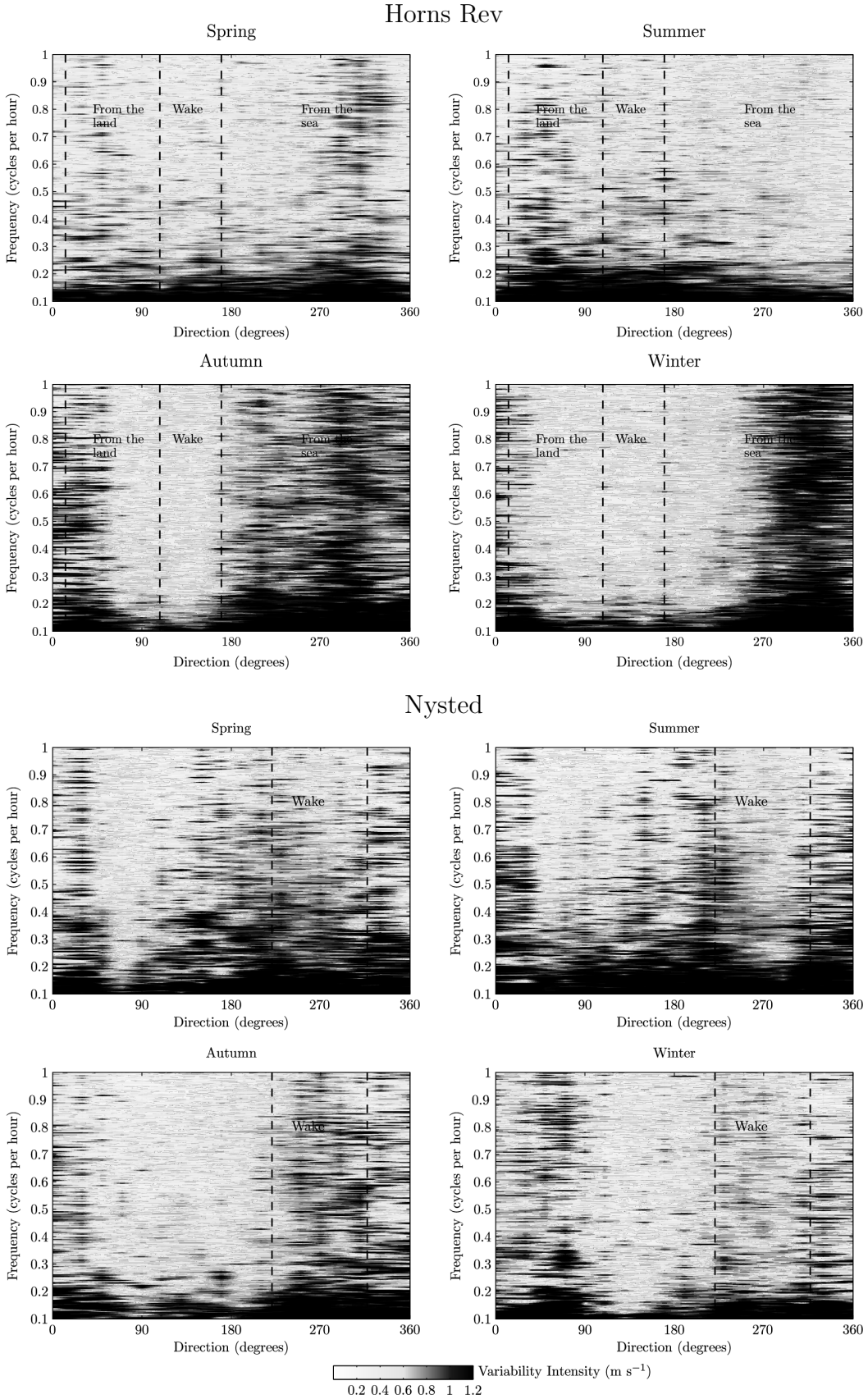


Figure 4.6: The Hilbert spectra of wind speed at 62 m at Horns Rev M2 (top four plots) and wind speed at 69 m at Nysted M3 (bottom four plots), conditionally averaged according to the season and wind direction. The vertical dotted lines indicate the range of directions that could be affected by the wake of the wind farm. For Horns Rev, wind directions corresponding to flow from the land and flow from the sea are also indicated.

4.3.3 Wind variability as a function of wind speed

Hilbert spectra as a function of wind speed for the two sites are shown in figure 4.7. Wind speed bins of width 2 m s^{-1} were used at both sites. The Horns Rev spectrum contains wind speeds up to 24 m s^{-1}, while the Nysted spectrum contains wind speeds only up to 20 m s^{-1}, since there were less occurrences of very strong winds at Nysted.

Both spectra show that higher average variability is observed for stronger wind speeds, although it should be noted here that higher wind speeds represent the tails of the distribution, so that there are progressively fewer observations in each bin as the wind speed increases. Many incidences of high wind speeds may occur as the local peak of a fluctuation, which will be associated with high variability by definition. In contrast, the trough of the fluctuation will be averaged with many more observations from different kinds of wind regimes.

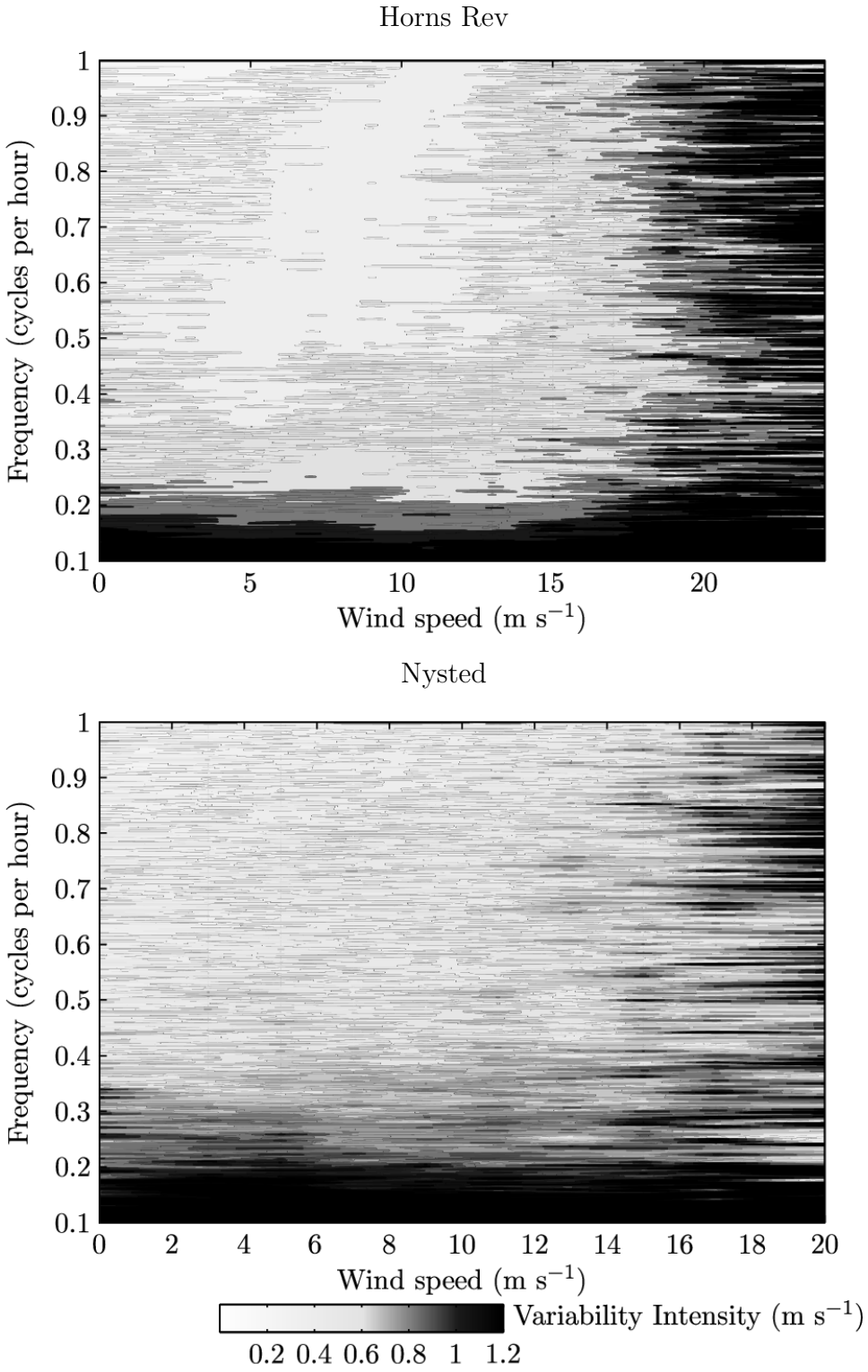


Figure 4.7: The Hilbert spectrum of wind speed for Horns Rev (top) and Nysted (bottom), conditionally averaged according to the wind speed

4.3.4 Wind variability as a function of pressure tendency

The Hilbert spectrum conditioned on the smoothed three hourly pressure tendency for the two sites is shown in figure 4.8. The pressure tendency bins have a width of 1 hPa per 3 hours. The pressure tendency here is used as a simple descriptor of the synoptic cycle, where rising pressure (positive tendency) generally indicates post-frontal conditions, and falling pressure (negative tendency) generally indicates pre-frontal conditions. More generally, a large negative or positive pressure tendency means that there are strong winds and dynamic weather conditions, probably in the vicinity of a front.

For both sites, wind variability is greater when the pressure is increasing quickly, and for Horns Rev it is also somewhat enhanced when the pressure is decreasing quickly. Interestingly, there is no increase in variability for rapidly decreasing pressure at Nysted. This difference between the two sites could reflect the different physical processes that are driving the mesoscale variability - for example, the North Sea is exposed to dynamic weather systems approaching from the west-northwesterly directions, while in the Baltic Sea, gravity waves under stable conditions have been reported (Larsén and Larsen, 2010).

Pressure tendency is an over-simplified metric of the synoptic cycle. The relationship between severe wind variability and the synoptic type at Horns Rev will be explored in more detail in chapter 5, where the observed wind variability at Horns Rev is related to the large scale weather patterns, wind classes and satellite images.

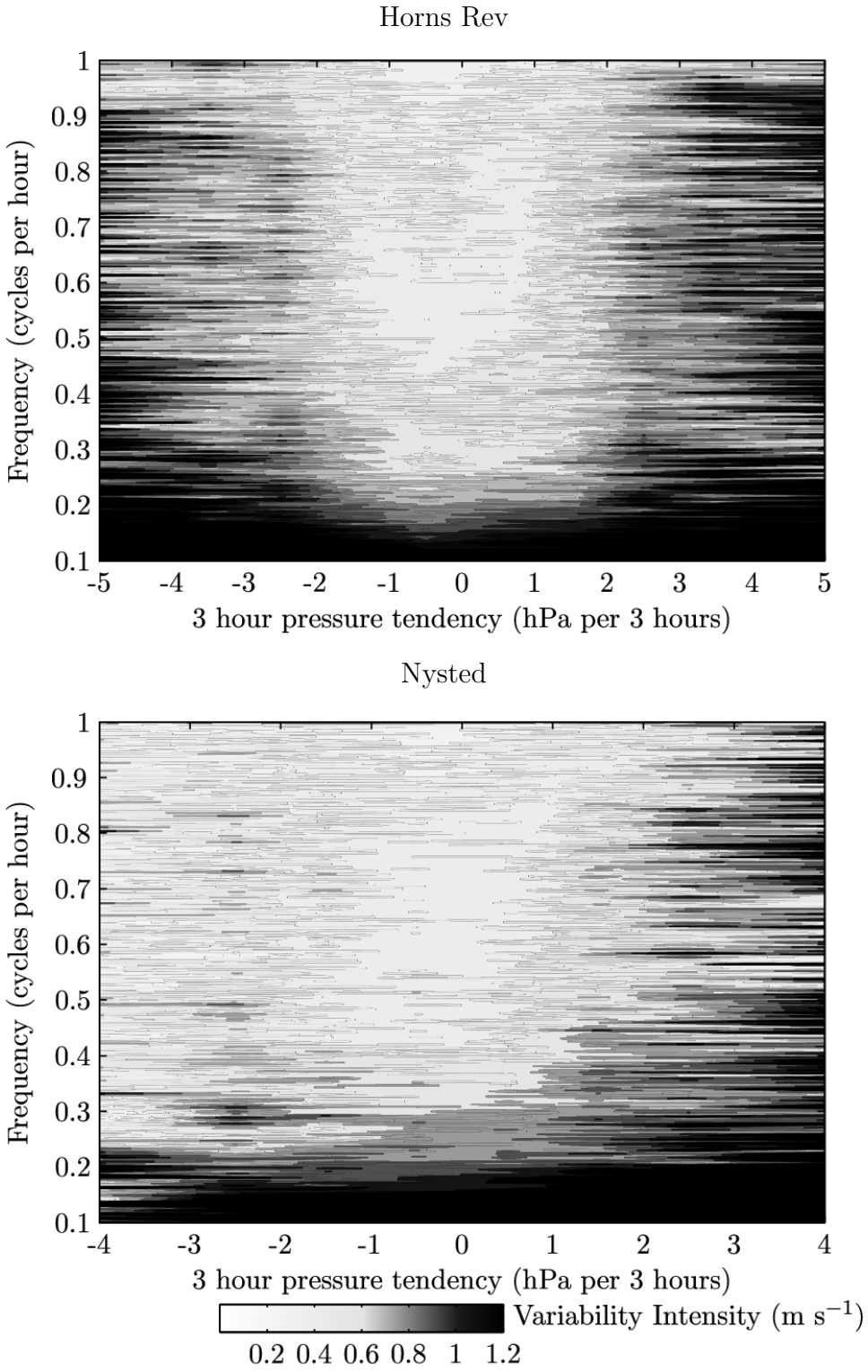


Figure 4.8: The Hilbert spectra for Horns Rev (top) and Nysted (bottom), conditionally averaged according to the smoothed three hourly pressure tendency. Positive values indicate rising pressure, while negative values indicate falling pressure

4.3.5 Wind variability as a function of air-sea temperature difference

Data for the air-sea temperature difference was only available for Horns Rev. The air temperature at 13 m and the sea water temperature at a depth of 4 metres were used to calculate the Hilbert spectrum as a function of air-sea temperature difference (figure 4.10). The temperature difference bins had a width of 1°C. The sea water temperature at a depth of 4 metres has been shown to be within about 0.6°C of the sea surface temperature measured by satellite at Horns Rev (Peña et al., 2008), so even though the sea water temperature is not exactly the same as the sea surface temperature, it is considered a good approximation.

The spectrum shows some trend for higher average variability for large negative air-sea temperature differences - that is, when the sea water temperature at 4 metres is more than about 2.5 degrees warmer than the air temperature at 13 metres. There also is also a slight increase in average variability when the sea water temperature is much colder than the air temperature. In other words, higher mesoscale variability is observed when the atmosphere is very thermally unstable, and somewhat also when it is very thermally stable. The occurrence of mesoscale variability when cold air is advected over the North Sea will be further discussed in chapter 5. Possible mechanisms for mesoscale wind variability in unstable conditions include cellular convection and convective rolls, while possible mechanisms in stable conditions include gravity waves. Furthermore, there could be a different contribution from the wake of the wind farm in stable and unstable conditions.

Like the results for pressure tendency, this result suffers from the fact that there are fewer and fewer observations as the bins move towards the tails of the air-sea temperature difference distribution. This explains the smooth appearance of the graph for air-sea temperature differences close to zero, and the streaky appearance of the graph as the magnitude of the air-sea temperature differences increases. This streaky appearance of the graph partially obscures the results.

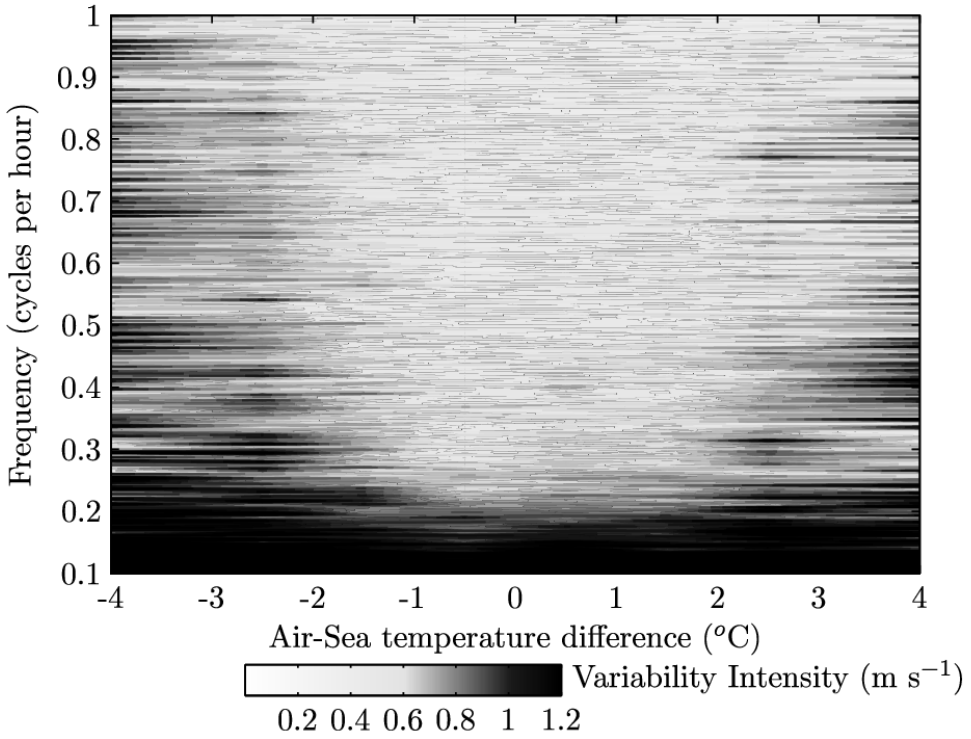


Figure 4.9: The Hilbert spectrum of wind speed at height 62 m at Horns Rev, conditionally averaged according to the air-sea temperature difference

4.3.6 Wind variability as a function of observed precipitation

There are several reasons to believe that wind variability is linked with precipitation. First, the regular patterns of rain showers seen on rain radars show that a location may be subjected to a nearly regular sequence of passing showers for some hours. As each rain shower passes, the gustiness associated with the outflow of the convective cell should create a periodic fluctuation in the wind speed. Further, according to Weusthoff and Hauf (2008), a distinct precipitation 'track', consisting of a collection of forming, splitting and dissipating precipitation cells under post-frontal conditions over land, can have life times of 29–40 min so that the changes in wind speed with the passage of showers should also follow these or slower time scales. The results already shown in this chapter also suggest a link with rainfall. Rain showers tend to approach the west coast of Denmark from the North Sea, and the rainiest months in western Denmark are August to December (http://www.dmi.dk/dmi/index/danmark/klimanormaler.htm).

There was insufficient data to calculate the Hilbert spectrum of wind speed at Horns Rev conditionally averaged on precipitation, particularly since the number of observations diminishes rapidly as the rain-rate increases, and occurrences of high rain-rate are actually the most interesting to analyse. A more simplified analysis was achieved by creating a scatter plot of the total amplitude of variability on time scales of 1–3 hours (as presented in figure 4.2) against the observed rain-rate. 'Rain rate' was a scalar time series, defined at time t as the total rainfall observed within a 90-min window before and after time t. This definition is justified by the idea that rain showers can be used as a proxy for generally convective conditions which are likely to last longer than the briefly observed passing rain shower.

The scatter plot of variability against rain-rate is given in Figure 9. The points in the scatter plot are aligned on particular rainfall values because precipitation is measured in discrete 0.25 mm quantiles. Further, the graph is hard to read because there are fewer and fewer values for increasing rain amounts. To aid interpretation, shading that shows the probability density function of variability for each given rain-rate have been added. From the shading, it is seen that with increasing rainfall, there tends to be an increase in the observed wind variability. There are also many high variability events when rain is not observed. Unfortunately an observation of no rain at a single point does not mean that there are not nearby rain showers, or that conditions of cellular convection and precipitation are not occurring. To draw a stronger relationship between rain showers and wind variability, spatial observations of precipitation (for example, from a rain radar) would be ideal.

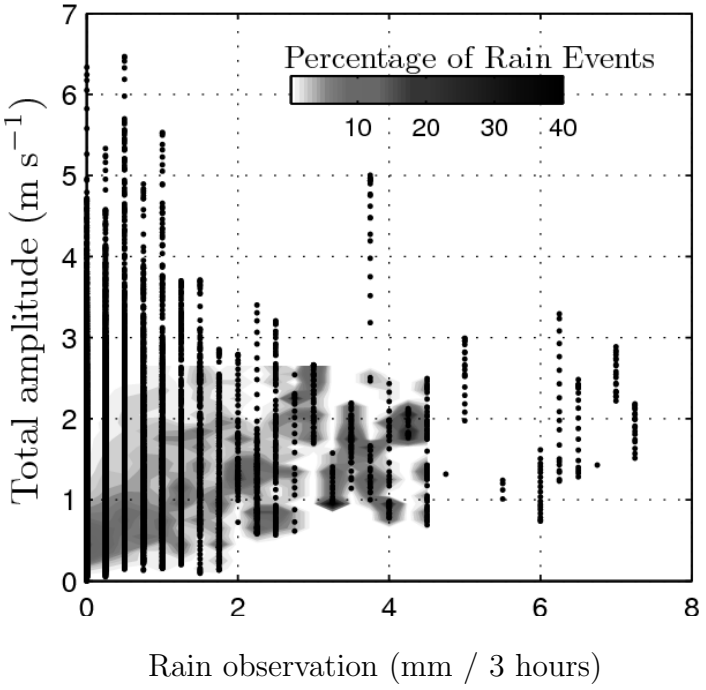


Figure 4.10: Scatter plot of variability on time scales of 1–3 hours at Horns Rev against observed rain-rate (points). Shading shows the probability density function of 1–3 hour variability for each rain-rate.

4.3.7 Further results regarding wind variability and the locally observed meteorological conditions

Further results about the observed wind variability at Horns Rev and the use of the Hilbert Huang transform as presented and discussed in Vincent et al. (2010a) and Vincent et al. (2010b). These results are summarised briefly here.

At Horns Rev, sonic anemometer wind speed measurements were available for a 10 month period from January–October 2004. This time series was analysed using the Hilbert-Huang transform in the same way as the 10 minute time series, such that direction dependent

spectra covering periods of 1–60 minutes could be calculated. The results, presented in Vincent et al. (2010b), showed that the trend of enhanced wind fluctuations in flow from the sea in autumn continued for all time periods within this frequency range, and the that trend of increased variability in winter for north-northwesterly wind directions also continued, but with less intense variability. There was also increased variability in the northeasterly sector during winter. As for the longer time periods, there was generally less variability in spring and summer than in autumn and winter.

An interesting outcome of these results was increased variability around the edge of the wake region, particularly during spring and autumn, for periods shorter than about 4 minutes. This result is consistent with results of Larsen et al. (2009), who showed that the wind speed in the wake of an 80-m diameter turbine rotor contains enhanced variance for frequencies greater than about 0.004 Hz.

Apart from wind direction, where there was a reasonable number of observations in each direction bin, the 10 month time series was not long enough to create conditional spectra for the other parameters that were considered for the 10 minute time series.

The marginal spectrum is the time integrated Hilbert spectrum, which as presented in Huang et al. (1998), is the average amplitude as a function of frequency in the same form as a Fourier spectrum. Marginal spectra were calculated for 3 different classes of stability, where stability was calculated using the Obhukov length following the methodology of Grachev and Fairall (1997). The results, which were compiled from a combination of cup anemometer data for 2000–2003 and sonic anemometer data for 10 months in 2004, showed that the spectral gap was well defined for stable conditions, but nearly non-existent for unstable and near-neutral conditions. This result was consistent with results of Gjerstad et al. (1995), who analysed maritime wind speed observations and found an obvious spectra gap for stable conditions, but no discernible spectral gap for unstable conditions.

In Vincent et al. (2010a), Hilbert spectra conditional on time of day for the 4 year time series of 10 minute cup anemometer observations were presented. While these results effectively demonstrated the conditional spectra methodology, and some weak trends between diurnal cycle and 1–3 hour variability were found, this result was not as significant or important as the later results in Vincent et al. (2010b).

4.4 Discussion and Conclusions

Analysis of time series of wind speed data using the HHT revealed a set of interesting patterns in the occurrence of large amplitude mesoscale wind fluctuations. The results suggested that the amplitude of mesoscale wind fluctuations at Horns Rev tends to be larger in the autumn and winter months, when the wind is from the sea, when the pressure tendency is rising or falling rapidly, when the wind speed is very strong, when the sea temperature is warmer than the air temperature, when precipitation is occurring and when the atmospheric conditions are unstable or near-neutral. These results were all calculated independently of one another, but they are almost certainly not independent - that is, several of these conditions are probably satisfied when large amplitude wind fluctuations are observed.

An obvious problem of analysing meteorological measurements from a measurement mast in

close proximity to a wind farm is that some measurements will be affected by the wind farm wake. However, there was no discernible signature of the wind farm wakes for either Horns Rev or Nysted in the analysis of 1–10 hour period variability. This suggests that although the wind farm wake contributes to turbulent fluctuations (Larsen et al., 2009) and has an effect on the mean wind (Christiansen and Hasager, 2005), it may not actually remove any energy from fluctuations on slower time scales. A useful aspect of the HHT is that the mean of the time series is subtracted at each step of the decomposition, so that the mean wind speed deficit in the wake should not have any influence on the spectrum.

The results presented in this chapter demonstrate that mesoscale wind fluctuations are not a constant feature of the wind speed, but depend strongly on the prevailing meteorological conditions. This highlights the fact that wind speed time series are non-stationary, which means that spectral analysis of a long time series will hide the time evolving component of the spectral behaviour. The well defined patterns between wind fluctuations and other meteorological parameters suggests that wind fluctuations could be predicted using statistical techniques such as multiple regression, or by combining auto-regressive models based on the variability time series with various explanatory variables.

This work was limited by the short length of the time series, especially for Nysted. However, longer time series are difficult to come by. In the future the shortfall in in-situ observations could be somewhat compensated by spatial variability information from satellite measurements, or from mesoscale modelling.

The differences between Nysted and Horns Rev, and between the different times of year and meteorological conditions suggest that some sites will be more susceptible to frequency mesoscale wind fluctuations than others. Sites that seldom experience large mesoscale wind fluctuations could potentially be more valuable sites for wind farms, since the problems associated with managing large wind fluctuations and sourcing the necessary reserve power will be diminished. This adds a new dimension to the problem of choosing the best sites for large offshore wind farms.

The methodology and results presented in this chapter and in chapter 3 relate to local measurements and locally observed wind fluctuations. Some of the results, such as the patterns between pressure tendency and wind variability, or between air-sea temperature difference and wind variability, will be the footprint of large scale weather conditions. The pattern between wind direction and wind variability is probably a combination of local effects (such as surface type and land use), and large scale effects (such as the tendency for mesoscale variability to occur in certain large scale flow conditions). In the following two chapters, the locally observed wind variability will be related to the large scale weather conditions using self-organising maps analysis of mean sea level pressure patterns and visual satellite pictures.

Chapter 5

Classification of severe wind variability days

5.1 Introduction

In chapter 4, it was shown that there are clear patterns that relate the local meteorological conditions to the amplitude of observed wind fluctuations. However, it is not only the local conditions that affect mesoscale variability, but the prevailing synoptic scale and mesoscale weather conditions. One reason why local observations are insufficient for characterising the parameter space within which mesoscale wind variability tends to occur is that the locally observed meteorological variables such as wind direction, stability or wind speed are modulated by mesoscale structures themselves, and as such will not always reflect the large scale weather conditions that dominate the situation. For example, the presence of a small low-pressure area embedded in north-westerly flow will result in locally observed wind directions that do not reflect the origin of the large scale flow.

In this chapter, three methods for classifying the occurrence of severe wind variability events are investigated, where severe wind variability events are considered as periods where the average 1–3 hour variability index (as defined in chapter 4) is in the top 5th percentile of all such periods. Firstly, the large scale weather patterns for an area extending from the southeast part of Greenland to northern Portugal in the north-south direction, and from Iceland to Poland in the east-west direction were classified. This domain took account of the fact that fronts and dynamic weather conditions often approach the North Sea from the westerly sector, and that the results from chapter 4 that suggest that many severe wind variability events occur when the wind comes from the westerly to northerly directions. Therefore the domain focuses more on the area to the west of Denmark than the area to the east. The large scale weather patterns are classified using Self Organising Maps (SOMs) (Kohonen, 2001).

Secondly, visual satellite pictures for each day in 2003 were classified according to the cloud types observed over the North Sea. For this classification, a subjective, visual classification

was used. Although this lacked the scientific rigour of an objective classification system, it still proved to be a useful categorisation. Finally, an established wind class classification system based on the wind speed, wind direction and stability in a large scale weather model, originally developed for use in developing numerical wind atlases was tested for its applicability for categorising mesoscale wind variability. Although the methodology uses values from a single point, it is not strictly a local classification system because it uses data from a large scale weather model which does not contain local effects anyway.

All three methods rely on the time series of wind variability for periods of 1–3 hours that was developed in chapter 4. Since all the methods are categorical classifications of the weather conditions at 24 hourly or 6 hourly intervals, the variability time series was also averaged to temporal resolutions of 6 and 24 hour respectively, with the averaging centered on the validity time of each classification. To convert the 24 hour variability time series to a categorical variable, 24 hour periods where the average variability was above the 95th percentile for all such periods were considered as 'extreme' events, and the number of extreme events in each category was evaluated. The choice of the 95th percentile as 'extreme' is arbitrary, since the real definition of 'extreme weather' comes from the point of view of those actually affected by the weather. However, for this work, the 95th percentile provides a useful working definition. In addition, the average meteorological conditions for each category in the synoptic and wind class classifications were calculated.

This chapter begins with a description of each of the three classification methods, and continues with a presentation of average meteorological conditions and the number of extreme variability days for each category of the three methods.

5.2 Synoptic Classification

Mean sea level pressure (MSLP) patterns for the North Sea region were classified using Self Organising Maps (SOMs) (Kohonen, 2001). SOMs are an automated pattern recognition technique, where the input data is mapped onto a low dimensional space (usually two dimensional) while preserving the topological relationships within the input data.

SOMs have been applied in a variety of different fields, including ecological sampling (Park et al., 2006) and bio-geochemical analysis of the Po river basin (Socal et al., 2008). In the field of synoptic classification, SOMs have been used by several authors, including Hewitson and Crane (2002), who classified large scale weather patterns over the north-east United States into 35 categories, and Michaelides et al. (2007), who presented 15-category and 20-category classifications of the synoptic conditions over Europe. Key results from the analysis of Michaelides et al. (2007) were that there was a high degree of persistence in synoptic type from one day to the next, that there was a strong annual pattern in synoptic types, and that these results did not fundamentally change when the number of categories was increased from 15 to 20.

SOMs are not the only way of making a synoptic classification. Indeed, many different methods have been applied to the problem. For example, several authors have applied principal component analysis (PCA) to weather patterns, then used a clustering technique on the component weights to create the classification (Cuell and Bonsal, 2009; Fealy and Sweeney, 2007; Romero et al., 1999). The advantage of these techniques is that the dimensionality of

the data is reduced prior to applying the clustering algorithm, and the relative weightings of different scales of atmospheric motion are taken into account. Casado et al. (2009) applied PCA to weather patterns over Europe, then classified individual time steps according to which principal component had the greatest loading. Despite the obvious difficulty of classifying days where there is a high loading on more than one principal component, they present a realistic set of 8 circulation types, based on the ERA40 500hPa geopotential heights for a 40 year period. A key observation from these results is that the North Sea region is predominantly subject to flow from the south-westerly or north-westerly directions. However, the analysis only extends slightly further north than Iceland, and as such does not illustrate the effect of the relative positions of low pressure systems in the North Atlantic Ocean.

A widely used classification for Europe is the *Grosswetterlagen* of the German Weather Service (DWD), a subjective scheme based on MSLP patterns and upper level steering flow that was recently reformulated in an objective framework by James (2007). An advantage of the method is that the categories are descriptive and related to intuitive weather scenarios (for example 'High Scandinavia-Iceland, Ridge C. Europe' or 'Low (Cut-Off) over Central Europe'). The Grosswetterlagen classification was used, for example, by Pryor and Barthelmie (2003) to connect an observed increase in winter time wind speeds over the Baltic region with an increase in westerly anticyclonic, westerly cyclonic and northwesterly cyclonic circulation types. They mention that such changes are important for the long-term value of wind energy resources over the region.

Philipp et al. (2010) presented an inter-comparison study of 23 different classification methods, including subjective classification types, threshold based methods, PCA based methods, lead algorithms, and optimization algorithms. SOMs are included in the last of these categories. Unfortunately, the SOMs classification was run for only 2000 iterations, and did not converge in this study. There was no clear favoured method, and it was concluded that although there were some differences between the methods, aspects of how the various methods are implemented are also have a strong influence on the results. The 23 classification schemes were used by Prudhomme and Genevier (2010) in the analysis of favourable circulation types for flooding in Europe. Using a metric that measured the spatial coherence of the favoured circulation types across flood measuring stations, they showed that the SOMs algorithm performed relatively poorly compared with the other algorithms.

There are no strong arguments for choosing one classification system over another. A few characteristics of the SOMs algorithm are actually appealing for this work. The first is that the topology of the distribution of patterns is preserved, which means that similar synoptic types will appear physically close to one another in the classification. This means that within the array of synoptic types, the transition from one type to another is obvious, and the set of synoptic types that are associated with severe wind variability are seen as a coherent group. Another advantage is that it is a clustering technique based on the data itself. No *a priori* decision about the final categories is required, since the process is initialised by a random choice of nodes. Finally, a SOMs classification is fundamentally different to a PCA decomposition, since the PCA decomposition results in a set of orthogonal basis functions. Each input vector can be reconstructed as a weighted sum of the basis functions, but the basis functions in themselves do not, in general, constitute plausible weather scenarios. A clustering or classification technique of some description must still be applied to the weights of the basis functions to create a classification system. In contrast, the SOMs algorithm is a clustering method applied directly to the input vectors themselves. A serious drawback of the SOMs method, which will be discussed further, is that an *a priori* decision must be

made about the number of categories.

5.2.1 The self-organising maps algorithm

The SOMs algorithm reduces the patterns in a set of N observations of a high dimensional variable to an array of $n \times m$ nodes. The $n \times m$ nodes are arranged so that similar nodes are close to each other in the array, while different nodes are far from one another.

Each of the N observations is considered as an input vector. In this case, an observation is one MSLP chart. The SOMs algorithm is then as follows: First, a set of initial nodes are chosen, either at random or as subsets of the data. The nodes are arranged in an array of dimension $n \times m$. The number of nodes is the number of categories that will be found in the data. Secondly, the first input vector is compared with each of the nodes, and the closest node to the data is selected and labelled as the 'favourite' node. The favourite node is then nudged towards the input vector. All other nodes within some radius R are also nudged towards the input vector, but with a decreasing effect depending on their proximity to the favourite node. The next input vector is then presented to the nodes, and the process repeated, until all input vectors have been presented to the nodes. This entire process is then repeated iteratively. The process is illustrated diagrammatically in figure 5.1.

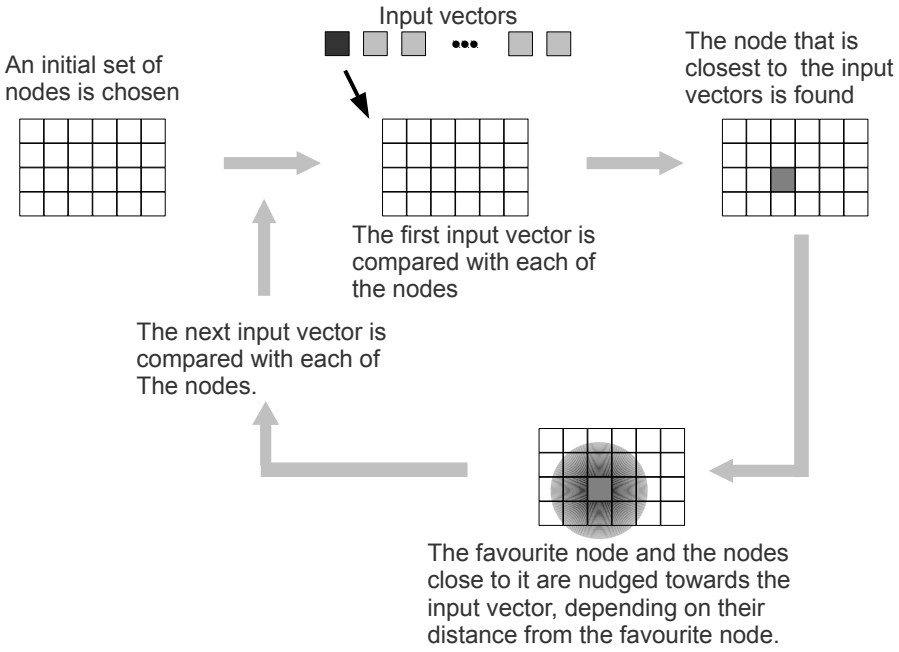


Figure 5.1: Schematic illustration of the SOMs algorithm

5.2.2 Methodology and Implementation

The SOMs classification of synoptic types was based on 1.5° resolution ERA interim analysis (Simmons et al., 2006/2007) MSLP data, for the period 1990–2009. Since the distance between the input vectors and the nodes is calculated as a euclidean distance, it makes

no difference whether the input vectors are considered in their actual matrix form, or as vectors. Therefore, for convenience, each two dimensional MSLP matrix was converted to a vector by concatenating its rows. One analysis per day, valid at 00UTC, was used. For each MSLP matrix, the mean pressure over the chosen domain was subtracted, so that the input vectors actually reflected the deviation from the mean pressure in the region, rather than the absolute pressure.

The SOMs algorithm was implemented using the R statistical software with the package 'Kohonen' (Wehrens and Buydens, 2007). In this implementation of the algorithm, the favourite node is updated by the new input vector using a weighted average of the node and the input vector. The weight of the input vector, or the 'learning rate' has a small value of 0.05 which decreases as the training proceeds. The units in the immediate neighbourhood are updated in the same way, and the size of the neighbourhood decreases during training, so that only the favourite node is updated after one third of the iterations are completed.

As already discussed, it is necessary to make an *a priori* choice about the number and shape of nodes in the SOMs array. Some authors have suggested algorithms for choosing the number of nodes. For example, Vesanto (2000) suggests that the number of SOMS, M can be found according to the formula $M = 5\sqrt{n}$, where n is the number of input vectors. This formula has been adopted by authors including Park et al. (2006). Other authors have tried different numbers of SOMS, then chosen the optimum number based on intuitive or physical arguments (Liu et al., 2006; Hewitson and Crane, 2002; Shahin et al., 2004). Giraudel and Lek (2001) give no justification for the number of nodes that they use.

Kohonen (2001) suggests using a Sammon map of the data to give a first guess of the ratio of the side lengths of the SOMs array. A Sammon map is a two dimensional projection of the data in such a way that the two dimensional Euclidean distance between each pair of points is as close as possible to the distance between them in higher dimensional space (Sammon, jr., 1969). If the distance between two points i and j in the higher dimensional space is D_{ij}, and the Euclidean distance between them in two dimensional space is d_{ij}, then the Sammon map is generated by minimising the quantity E,

$$E = \left(\sum_{i=1}^{N} \sum_{i<j} D_{ij} \right)^{-1} \left(\sum_{i=1}^{N} \sum_{i<j} \frac{(D_{ij} - d_{ij})^2}{D_{ij}} \right). \tag{5.1}$$

The Sammon map for four years of the input data is given in figure 5.2. Each dot represents one surface pressure chart, so two points that are close neighbours will represent two days with very similar large scale weather patterns. The Sammon map has no distinct clusters, indicating that the data are fairly continuous. The overall bounding box of the Sammon map is approximately square. Therefore, a square array of SOMs was chosen. The reason for this is that the SOMs array must span the same variance as the original data in both of the two available dimensions.

de Bodt et al. (2002) proposed assessing the number of nodes that should be used in a SOMs array using the ratio of the mean to the standard deviation of the total quantisation error, which was named the *coefficient of variation* (CV). The total quantisation error is defined as the sum of all the distances between the input vectors and their final best fitting nodes. The mean and standard deviation of the total quantisation error are found by bootstrap

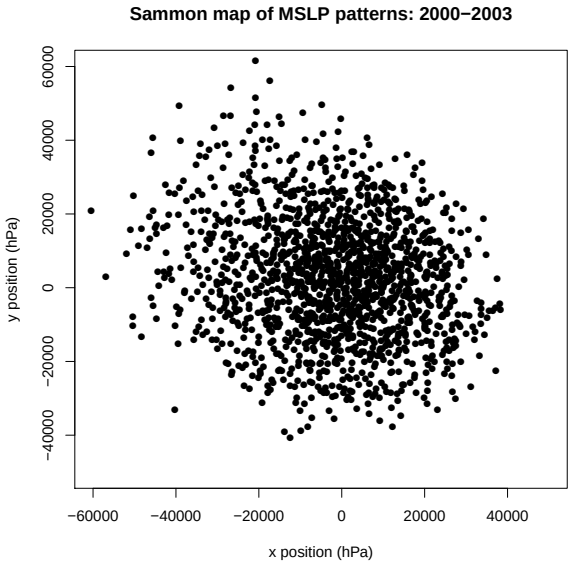


Figure 5.2: Sammon map of the input data. The units are in hPa, because the distance is the Euclidean distance between pairs of MSLP maps.

sampling the original set of input vectors N times (where 100 is a suggested value for N). The methodology of de Bodt et al. (2002) was tested, although it was necessarily computationally expensive, since it required 100 realisations of the SOMs analysis for each possible dimension of the SOMs array. Due to this computational demand, the data set was restricted to the first 1000 (about 3 years) input vectors. For this subset of input vectors, a 6 $\times$ 6 array gave the minimum value of CV. However, repeating the exercise for different numbers of input vectors seemed to suggest that the recommended size of the array increased as the number of input vectors increased. This is therefore different to our real aim, which is to show whether is an intrinsic or natural division of the data that is robust to the number of input vectors.

The SOMs classification was conducted for square SOMs arrays with side lengths from 2 to 10 nodes, to test for the best size of the array. The SOM arrays of dimension 2 $\times$ 2, 4 $\times$ 4 and 6 $\times$ 6 are shown in figures 5.3, 5.4 and 5.5 respectively. It is seen that there are some common themes between the arrays, namely that the flow over the North Sea is largely modulated by the relative positions of the high pressure ridge and the low pressure trough or circulation over the region. The percentage of days in each category for the 6 $\times$ 6 SOMs array during the study period 2000–2003 is shown in figure 5.6, which demonstrates that the method does not impose any condition that there must be an equal number of cases in each category.

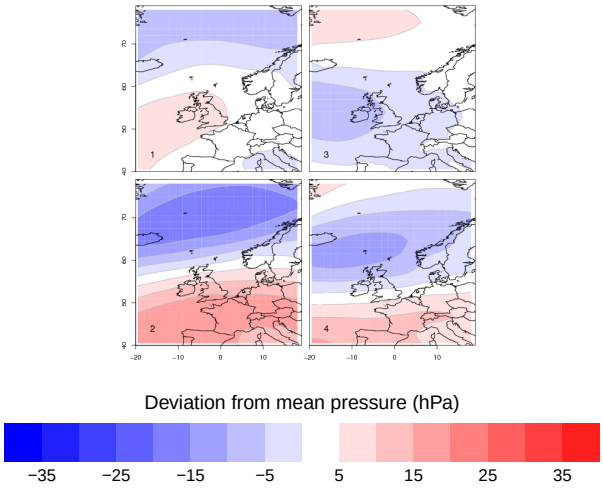


Figure 5.3: 2 × 2 SOMs array

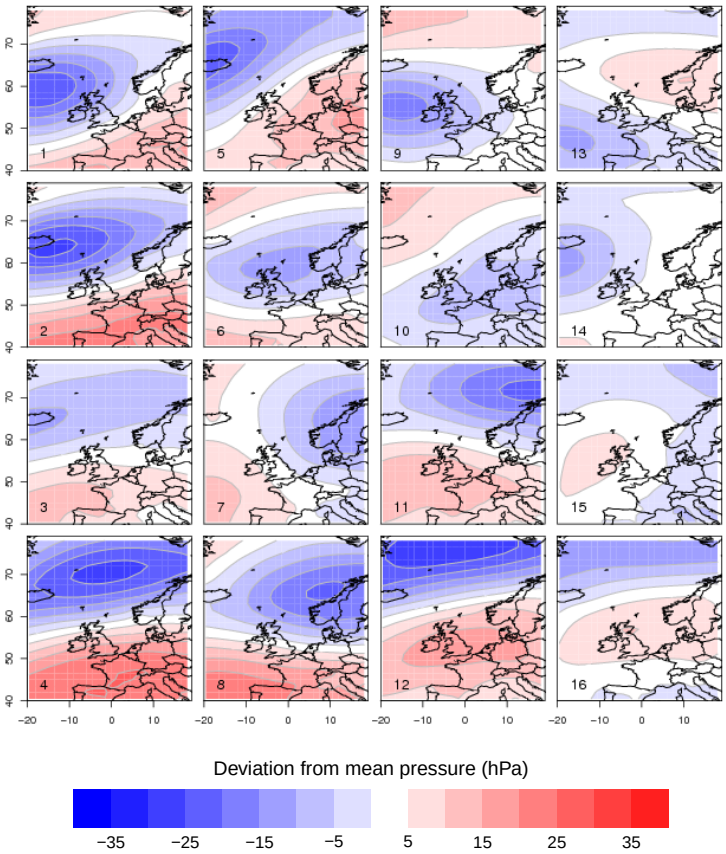


Figure 5.4: 4 × 4 SOMs array

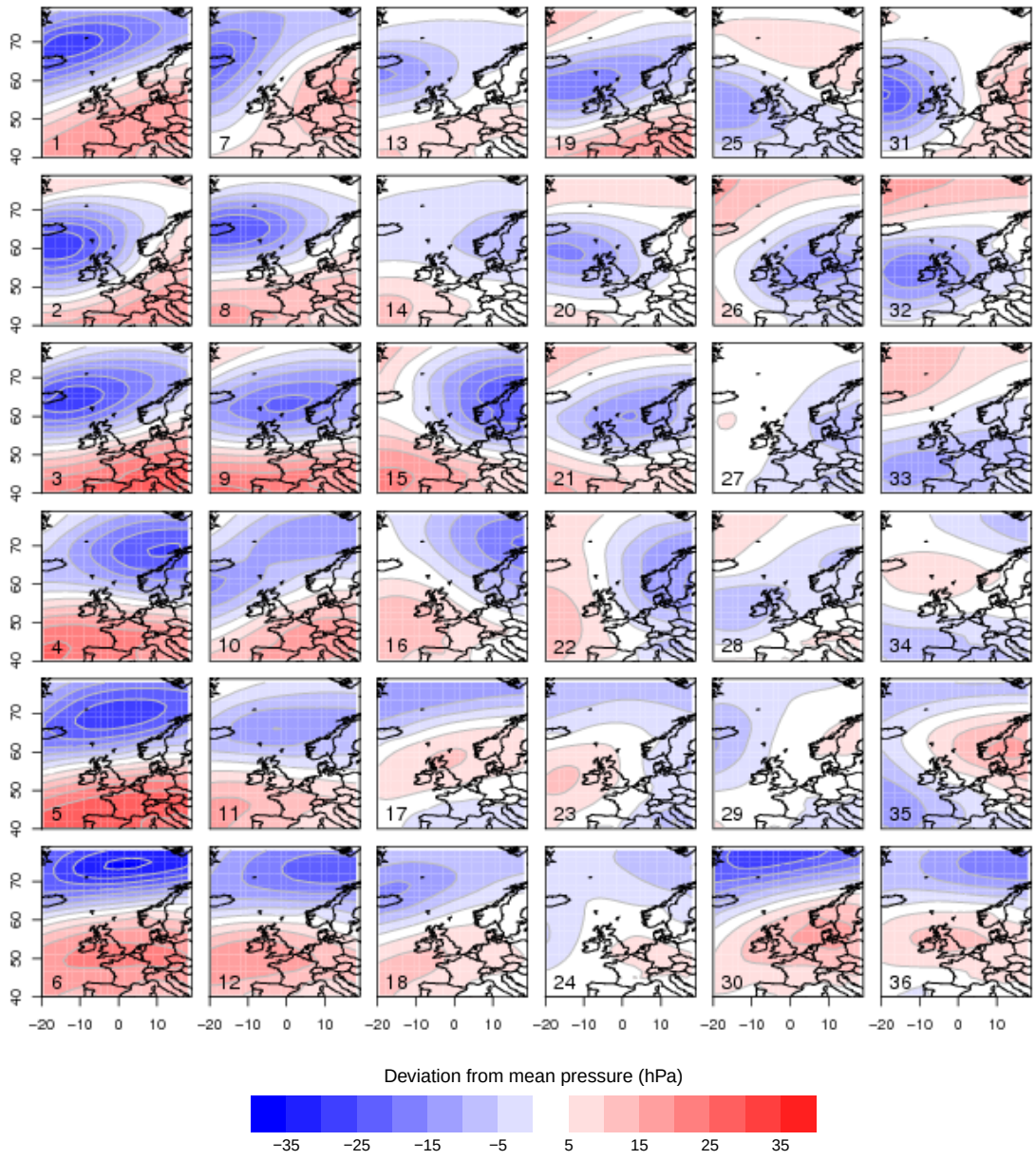


Figure 5.5: 6 × 6 SOMs array

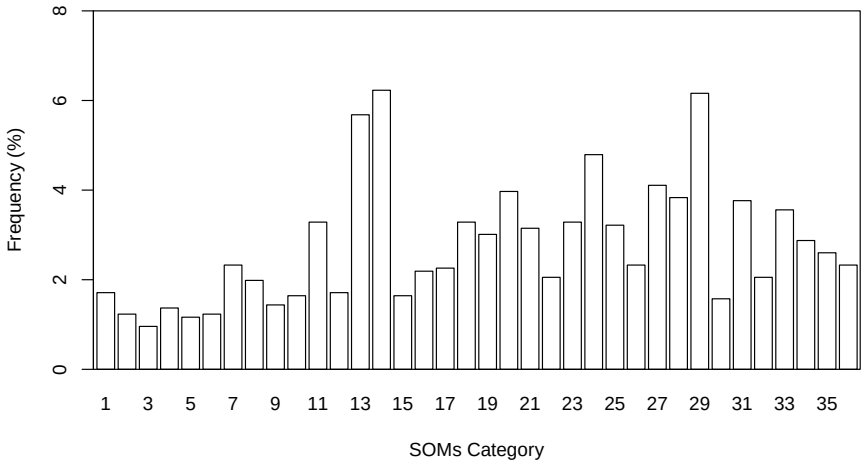


Figure 5.6: The percentage of days in each 6 × 6 SOMs category in the period 2000-2003

Since the SOMs nodes should retain the topology of the original data set, the input vectors within particular SOMs categories should be grouped close to one another on the Sammon map (figure 5.2). In figure 5.7, sections of the Sammon map in figure 5.2 are reproduced, with colour-coding for the different categories of the 6 × 6 SOMs array. The three plots relate to the top left of the SOMs array, the centre and the bottom right. The pressure patterns that lie in the top left corner of the SOMs array, where the weather over the North Sea is dominated by a low pressure system with a centre near Iceland, occupy a particular region of the Sammon map. Similarly, the bottom left corner of the SOMs array, where the weather over the North Sea is dominated by a high pressure ridge, are grouped closely on the Sammon map. On the other hand, the pressure patterns that fall in the centre of the SOMs are fairly dispersed on the Sammon map, which might be because they represent transitional situations. It should also be noted that both the SOMs array and the Sammon map can only be approximations to the higher dimensional topology of the data.

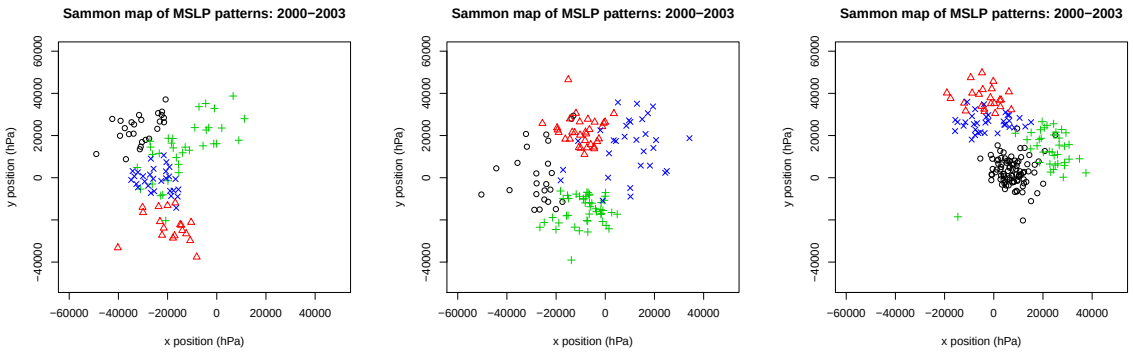


Figure 5.7: Sections of the Sammon map showing groups of nearby SOMs categories, where each input vector is coded by colour and symbol according to its SOMs category. Left: SOMs categories 1 (o), 7 (+), 2 (△) and 8 (×). Centre: SOMs categories 15 (o), 21 (+), 16 (△) and 22 (×). Right: SOMs categories 29 (o), 35 (+), 30 (△) and 36 (×). The units are in hPa, because the distance is the Euclidean distance between pairs of MSLP maps.

5.3 Satellite Classification

Satellite pictures were obtained from the Terra and Aqua MODIS satellites, which are polar-orbiting satellites travelling around the earth in opposite directions to one another. Since visual satellite pictures were being used, the analysis was limited to daylight hours. Fortunately, one of the two satellites passed over the North Sea on nearly every day of 2003 within about 2 hours of 1200 UTC, which meant that visual images were available even during winter. Images were downloaded from the LAADS Web of the NASA Goddard Space Flight Center (http://ladsweb.nascom.nasa.gov/), and processed using the software HDFLook.

Automatic pattern recognition algorithms for classification of clouds and air masses from satellite pictures have existed for several decades. For example, Pankiewicz (1997) developed a neural network classification system for METEOSAT images based on pre-identified subjective air-mass classifications and the spectral, textural, spatial and contextual properties of the pictures. Addressing a similar problem in ground based whole-sky images, Heinle et al. (2010) developed a cloud classification system using k-means neural networks, which is a similar system to the SOMs classification.

Despite these interesting applications of pattern recognition algorithms for cloud classification, a subjective classification method was chosen here. This was due both to the small data set (one image per day for 365 days), and due to the flexibility and simplicity of the method. The area under consideration is the North Sea between Denmark and England in the east-west direction, and between the south of England and the southern most tip of Norway in the north-south direction.

A preliminary inspection of the 365 satellite pictures was used to define a subjective set of 16 categories of cloud patterns, supplemented by a category for pictures that were unclear or impossible to classify, and a category for pictures that did not show a sufficiently large area of the North Sea to make a meaningful classification. This gave a total of 18 categories, which are listed below.

Category	Description
1	Open cellular convection, whole North Sea area
2	Open cellular convection, part of the North Sea area
3	Confused skies, with possible areas of open cellular convection
4	Convective rolls or possible convective rolls, combined with areas of open cellular convection
5	Convective rolls or possible convective rolls, combined with other cloud types
6	Total cloud cover / nearly total cloud cover - hard to see any structure
7	Clear skies / nearly clear skies
8	Front or well defined cloud band dominating the region
9	Isolated large convection
10	Scattered large convection
11	Closed cellular convection
12	Areas of small cumulus / fair weather cumulus (could be mixed with areas of clear skies)
13	Mainly areas of stratiform cloud, possibly mixed with ares of clear skies
14	High cloud dominating the region - hard to see any structure beneath
15	Areas of stratocumulus (could be mixed with areas of clear skies)
16	Confused skies, unknown classification
17	Insufficient area shown in satellite picture
18	Clear low pressure circulation in the North Sea

Table 5.1: Categories for the satellite classification.

Although the classification was subjective, the images were classified in a random order without reference to time of year, meteorological conditions or observed variability.

5.4 Classification by wind classes

As already discussed, the SOMs analysis is a large scale analysis that does not take the local conditions into account. An alternative classification system that has been used effectively are the wind classes used in numerical wind atlases (Badger et al., 2007). The wind classes are calculated by taking the geostrophic wind direction, geostrophic wind speed and stability into account, where the stability is quantified by the inverse Froude number. The data for calculating the wind classes typically comes from a single point in a large scale model, in this case the NCEP/NCAR reanalysis. Thus it is a local classification system, but only in the sense that it relies on model data at a single point. Mesoscale effects will not affect the classification, because the data does not come from a mesoscale model. The wind classes are calculated so that there are a similar number of data in each category.

An existing 135 category classification of wind classes for a point near the Horns Rev wind farm, which was used in Badger et al. (2007) and generously made available through personal communication with the authors, was analysed in the same framework as the SOMs analysis.

5.5 Results

5.5.1 Synoptic Classification

The number of severe variability days in each SOMs category for each size of SOMs array was determined - as an example, the results for the 2 × 2, 4 × 4 and 6 × 6 arrays are shown in figure 5.8. The results from the SOMs analysis with different sized arrays are consistent with one another, and show that the categories where the most variability is observed are mainly those where the geostrophic wind direction over the North Sea is in the west to northwesterly sector, or where there is a low pressure centre in close proximity to the North Sea itself. The 2 × 2 SOMs array is clearly under-resolved, since it does not differentiate between southwesterly and northwesterly flow. Arrays larger than 6 × 6 were difficult to analyse because they had a lot of very similar categories. The 6 × 6 SOMs map was chosen for subsequent analysis based on the heuristic argument that the synoptic patterns where a large number of extreme variability days occur are still obviously different from one another. The frequency of days falling in each 6 × 6 SOMs category for the period 2000–2003 is shown in figure 5.6. Otherwise, no attempt to maximise the number of extreme events by choosing the number of SOMs was made.

For the 6 × 6 SOMs array, category 15 stands out as the category with the most extreme variability events, with 15 cases. Since there were 24 category 15 days in the period 2000-2003, this means that severe wind variability was observed on 63% of category 15 days. Category 21 has 9 cases, categories 11, 16 and 22 have 6 cases, and category 4 has 5 cases. Altogether, 47 out of the 73 extreme variability days fell within these 5 SOMs categories. There are 19 categories that do not have any extreme variability cases.

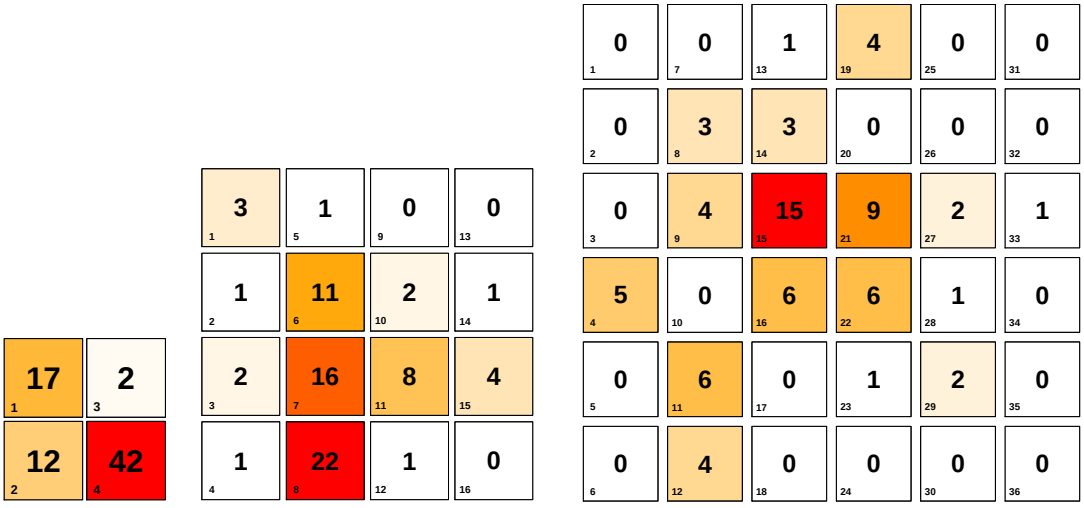


Figure 5.8: The number of extreme variability days in each SOMs category for the 2 × 2 array (left), the 4 × 4 array (centre) and the 6 × 6 array (right). The arrays have the same configuration as the SOMs arrays in figures 5.3, 5.4 and 5.5, and the SOMs index is indicated in the bottom left of each panel. Colour shading corresponds to the number of extreme variability days in each category, where the categories with the lowest number of extremes are white and the category with the highest number of extremes is red.

The MSLP patterns for the period 2000–2003 were classified in to SOMs categories for each

6 hour period. For each of the 36 SOMs categories, the average values for average variability, maximum variability, wind speed, wind direction, temperature and pressure for Horns Rev M2 were calculated, based on 6 hours centered on the validity time of the MSLP chart. The average wind direction was calculated so that values between 0 and 90 degrees and values between 270 and 360 degrees were not erroneously averaged through simple averaging. This is referred to as the 'circular average'. Days when data was not available were neglected from the analysis. The results are shown in figure 5.9, where the SOMs categories have been sorted by the average daily variability. The results show that the SOMs classes with the highest average daily variability are 15,21,22,16,9,4 and 11. The bars, which show the 90th and 10th quantiles, show that while the data within each class do show considerable spread, the trends are clear. All categories contain at least 50 data points. For some of the other meteorological variables, such as wind direction and in some cases, temperature (eg. category 5), the range between the 10th and 90th quantiles is surprisingly small. This demonstrates that the relationship between the large scale flow and the locally observed meteorological variables is relatively strong.

Interestingly, the graph shows that the SOMs classes with the highest variability are not necessarily those with the strongest wind speed. For example, category 3 has the highest average daily wind speed, but a reasonably low average daily variability. The daily average wind direction shows that the classes with highest variability generally have wind directions in the northwesterly sector, but this range of directions is not unique to high variability days - for example, classes 36 and 24 have average wind directions of around 300 degrees, but low average variability. There is no particular trend in the temperature of the high variability days, but high variability tends to occur with low average daily pressure. These results are consistent with the results presented in chapter 4, where it was shown that the maximum amplitude of variability tends to occur in the west to northerly sector, and when the pressure is decreasing rapidly.

5.5.2 Satellite Classification

The number of severe variability days falling in each satellite classification category was calculated for the year 2003. As for the SOMs analysis, a severe variability day was defined as a day where the average 1–3 hour variability was in the top 5th percentile. In this analysis, the 24 hour average was centered around the satellite image time of 1200 UTC, unlike the SOMs analysis where the 24 hour average was centered on the analysis time of 0000 UTC. This difference was simply for the convenience of aligning the days with the satellite image times.

The number of extreme variability days in each satellite classification category is shown in figure 5.10, with the total number of cases in each category indicated in brackets above the bars. While the amount of data is too small to draw strong conclusions, the graph does show that 14 out of the 19 extreme variability days occurred in categories 1–4 - that is, in the four categories with definite or possible observed open cellular convection (OCC). Of the remaining five severe variability days, one occurred with total cloud cover, two occurred with scattered large convection, one occurred with areas of small cumulus and one occurred with confused skies or unknown classification. The number of extreme variability days could be increased by lowering the 95th percentile definition. On the other hand, figure 5.10 shows the satellite types of the most extreme cases. A longer time series of satellite pictures would

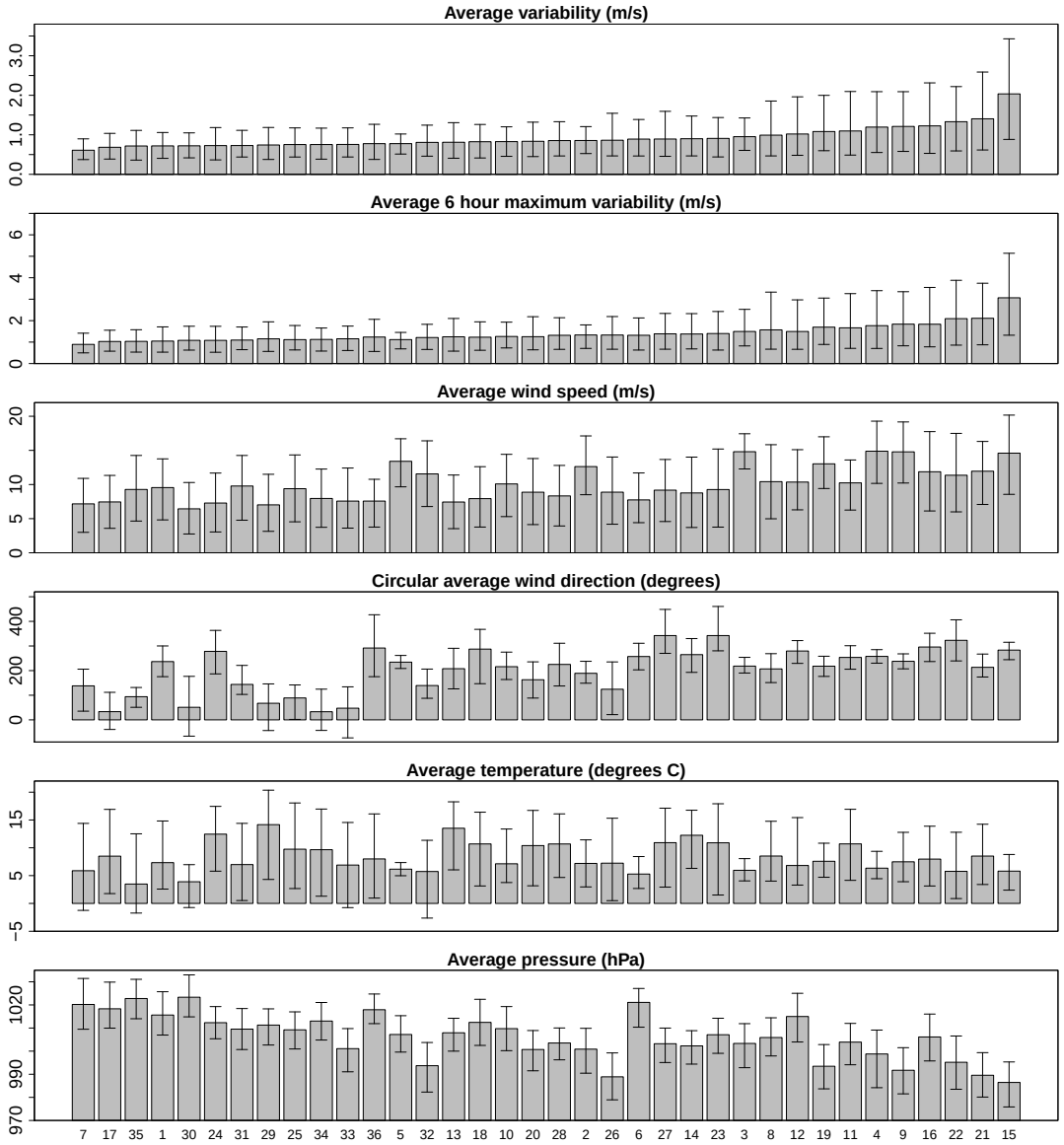


Figure 5.9: Average daily observed meteorological variables at Horns Rev for each SOMs category, ordered as a function of average daily variability. The SOMs categories of each bar indicated on the horizontal axis relate to the SOMs array in figure 5.5. The bars show the 10th and 90th quantiles. The axis on the average wind direction plot extends beyond the range 0–360 degrees to allow for meaningful quantiles.

also be highly beneficial to the analysis.

Satellite pictures for the eleven days in satellite categories 1 and 2 which were also extreme wind variability days are shown in figure 5.11. These are all clear cases of OCC over all or part of the North Sea. Furthermore, the sizes of the cells are qualitatively similar for all eleven cases.

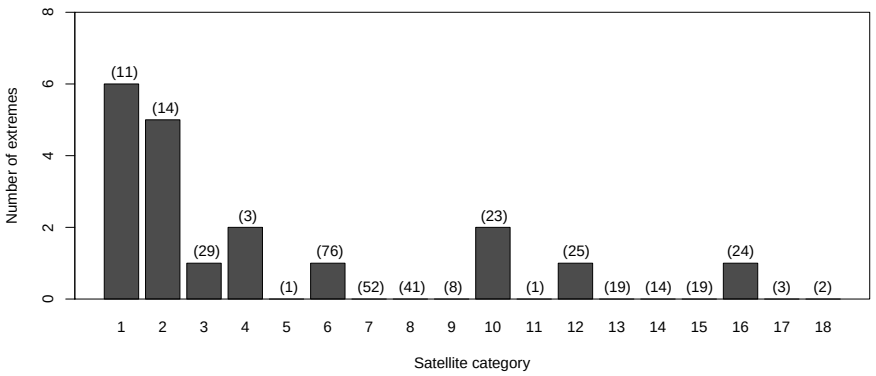


Figure 5.10: The number of extreme variability days in each of the satellite classification categories described in table 5.1. The total number of days in each category in 2003 is indicated in brackets above the bars.

5.5.3 Classification by wind classes

The number of extreme wind variability days in each wind class for the years 2000–2003 was calculated (figure 5.12), and the average meteorological variables for each wind class were calculated (figure 5.13). The patterns in wind speed and wind direction follow regular patterns because of the systematic way in which the wind classes were calculated for different groups of wind speeds and directions. Because of the large number of wind classes, some categories contained very few data points. Categories with less than 30 data points have been indicated with a dot on the average variability plot, where it is seen the these categories are concentrated around northerly to easterly wind directions and very strong wind speeds. The bars show the 90th and 10th quantiles.

The wind classes with more than three extreme variability days are described in table 5.2, where the wind classes are defined by the boundaries of their wind speed, wind direction and stability bins. Consistent with the SOMs analysis, the wind classes show that the most variability is found when the wind is coming from the west to northwesterly directions. In the wind classes which differentiate between unstable and stable flow (not all do, since the wind classes are designed to have the same number of observations in each class, and this subdivision is neglected if there is insufficient data), the wind classes with more than 3 extreme variability days are all unstable classes.

5.5.4 Typical wind speed time series observed in certain SOMs categories

The difference in average wind variability between the different SOMs categories indicates that the observed time series for particular SOMs categories should display noticeably different statistical patterns. Visual confirmation of this is given in figures 5.14 and 5.15, where 24 hour time series of wind speed at Horns Rev M2, centred at the SOMs analysis time of 00UTC, for all the days falling in SOMs categories 3 and 15 in the period 2000–2003 are shown. These two groups are exemplary of the fact that the strongest wind speeds are not

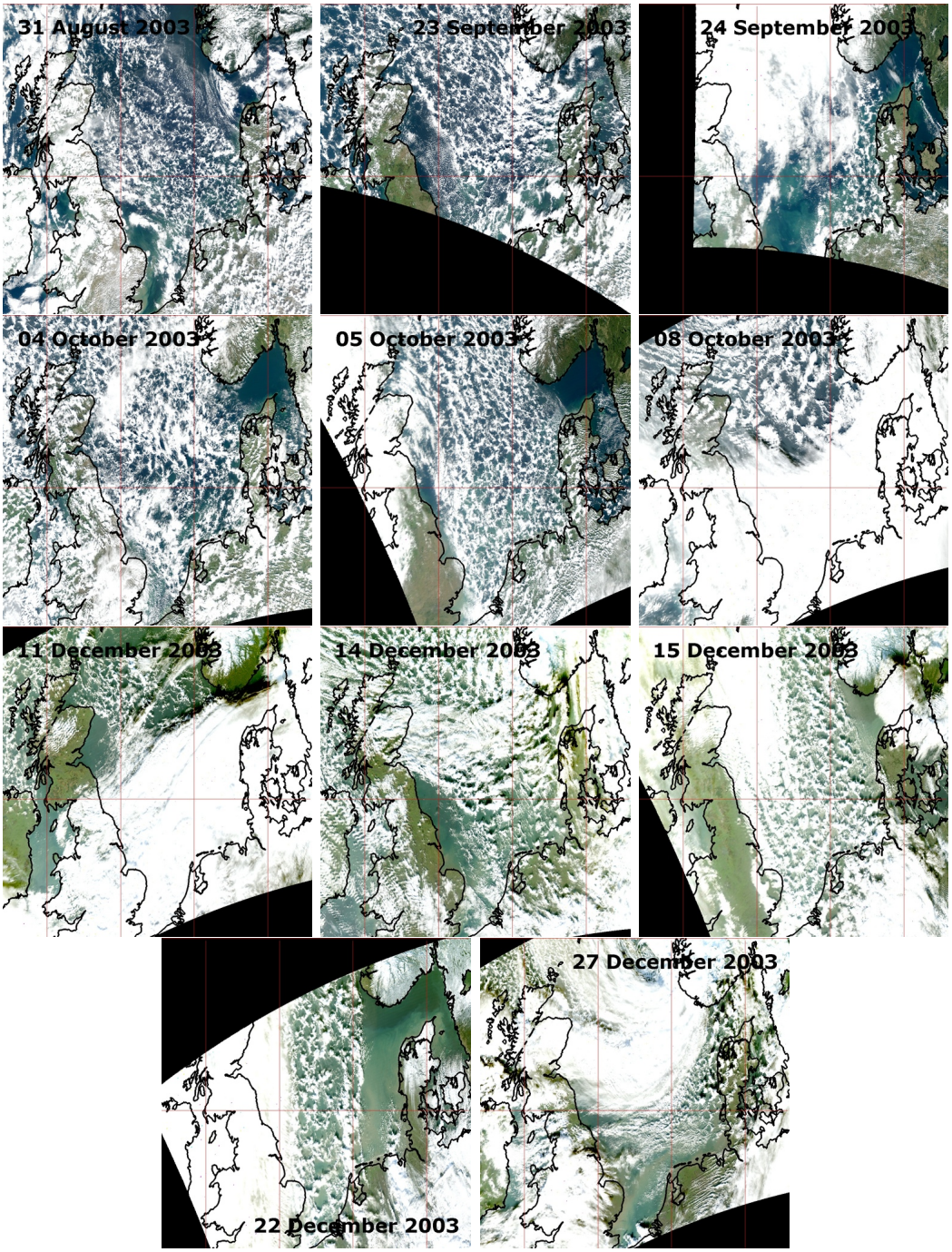


Figure 5.11: Visual Satellite images from the TERRA and AQUA MODIS satellites (sourced from http://ladsweb.nascom.nasa.gov) for satellite group 1 and 2 days which recorded extreme mesoscale wind variability. All satellite image are valid within two hours before or after 1200 UTC.

necessary related to the strongest wind variability. Group 3 has the highest average daily wind speed, but a relatively low average daily variability. In contrast, group 15 has a high average daily wind speed, and the highest average daily variability. Inspection of figures 5.14 and 5.15 show that the group 3 time series are smoother than those in group 15, where large wind fluctuations are observed on most days, often with amplitudes of up to about 10 m s^{-1}. Furthermore, for some of the days (such as 2001 days 360–361 and 2003 day 280), the fluctuations occur mainly within the bounds of 5 m s^{-1} and 15 m s^{-1} that correspond to the

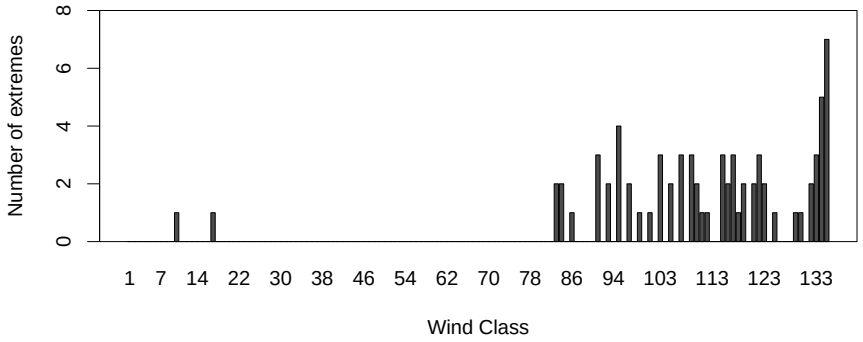


Figure 5.12: The number of extreme variability days observed at Horns Rev for each wind class category during the period 2000–2003

Wind class	Direction (min)	Direction (max)	Speed (min)	Speed (max)	Fr^{-1} (min)	Fr^{-1} (max)
91	225	255	5.32	8.17	-1000000	2
95	225	255	10.77	13.98	-1000000	1
103	255	285	5.62	8.57	-1000000	2
107	255	285	11.07	14.08	-1000000	1
109	255	285	16.14	18.90	-1000000	1000000
115	285	315	5.37	8.16	-1000000	2
117	285	315	8.16	10.72	-1000000	2
122	285	315	18.52	24.41	-1000000	1000000
133	315	345	14.60	17.15	-1000000	1000000
134	315	345	17.15	22.18	-1000000	1000000
135	315	345	22.18	90.00	-1000000	1000000

Table 5.2: Wind speed, wind direction and stability definitions for wind classes with more than three extreme variability days

most responsive part of the power curve, and are therefore of strong relevance to fluctuating power production from wind farms.

Since the satellite classification and the SOMs classification both showed categories with a strong likelihood of severe wind fluctuations, the two classification systems should also be consistent with one another. The number of days falling in each combination of the two systems is shown in table 5.3, where the grand total of the table is 365 days. The five SOMs category rows which had more than 5 extreme variability cases during the period 2000-2003 are shaded, and the four satellite category columns relating to OCC or possible OCC are also shaded.

Table 5.3 shows that OCC was observed on all 5 of the SOMs category 15 days, on 2 out of the 4 SOMs category 21 days, on 4 out of the 9 SOMs category 16 days and on 6 out of the 8 SOMs category 22 days. For SOMs category 11, OCC was observed on only 4 out of 16 days, although there were 5 days with total cloud cover which could have been obscuring other clouds underneath. Other SOMs patterns in which multiple definite cases of OCC was observed are 23, 27, 33 and 34.

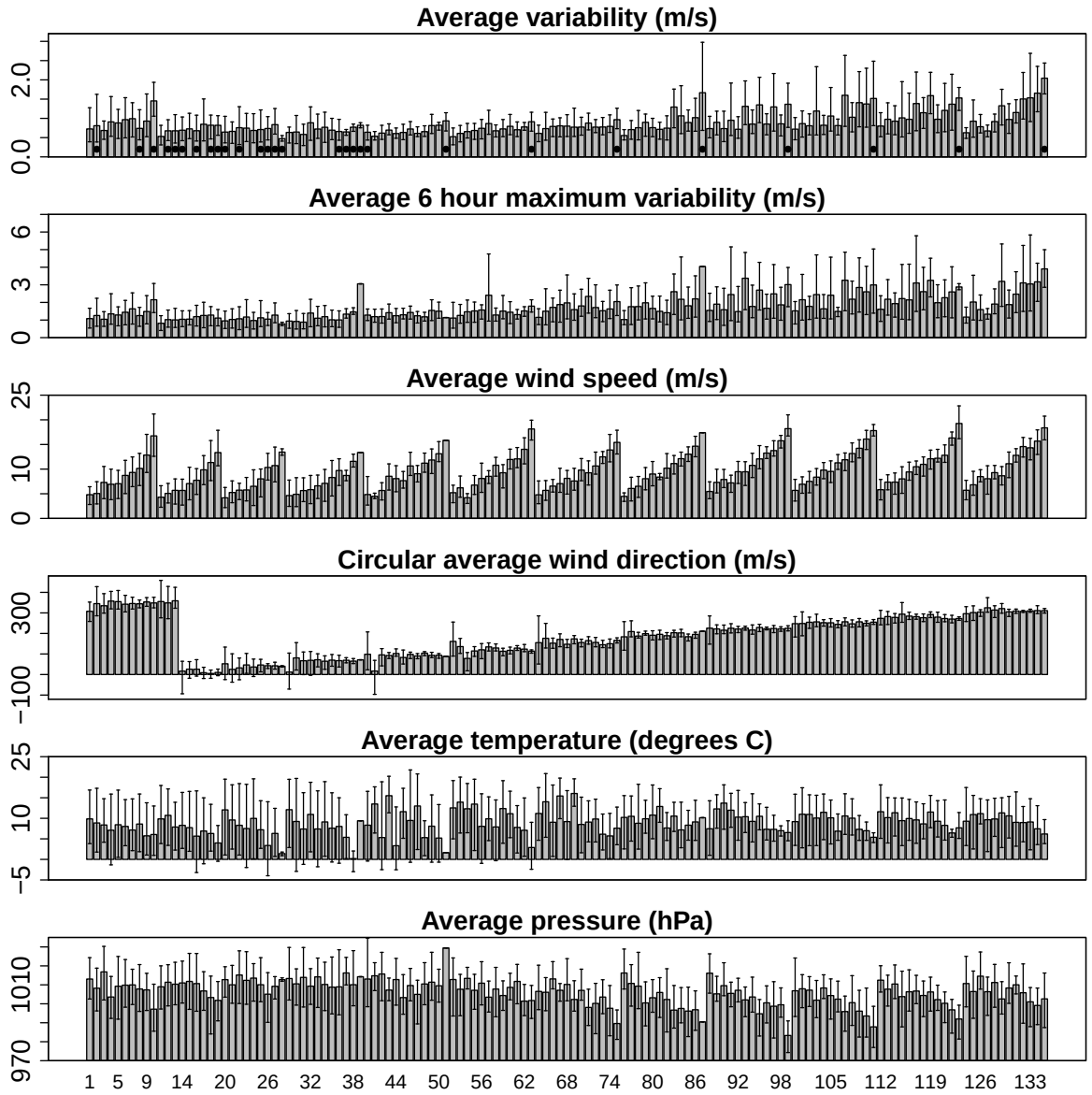


Figure 5.13: Average daily observed meteorological variables at Horns Rev for each wind class during the period 2000–2003

Although one year of data is not enough to classify days into so many combinations of categories, it is clear that the occurrence of OCC is strongly linked with certain synoptic patterns. The OCC is mainly related to those synoptic patterns with cold northwesterly flow, and for both classification systems, this is where many cases of severe mesoscale wind variability are experienced.

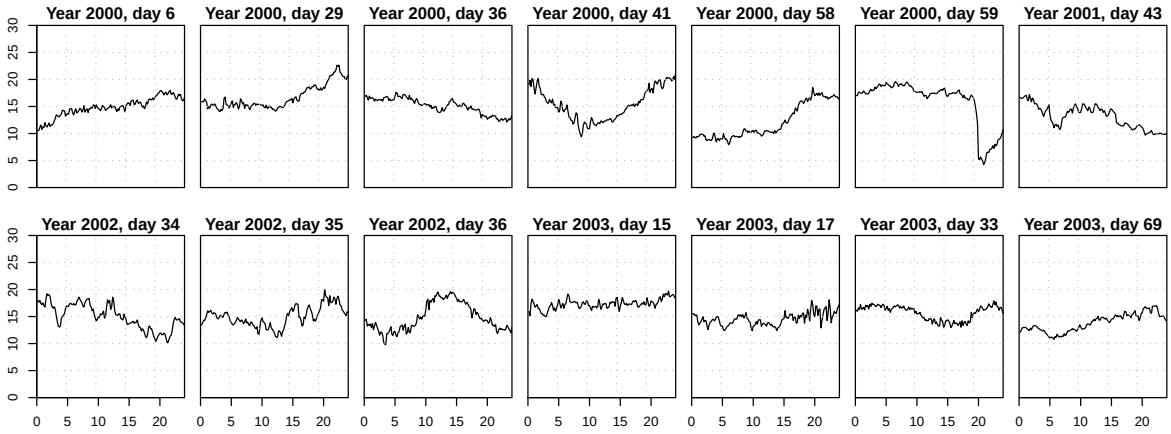


Figure 5.14: 24 hour time series of wind speed at Horns Rev M2 (centred at 00UTC) for the 14 category 3 cases that occurred during the period 2000–2003

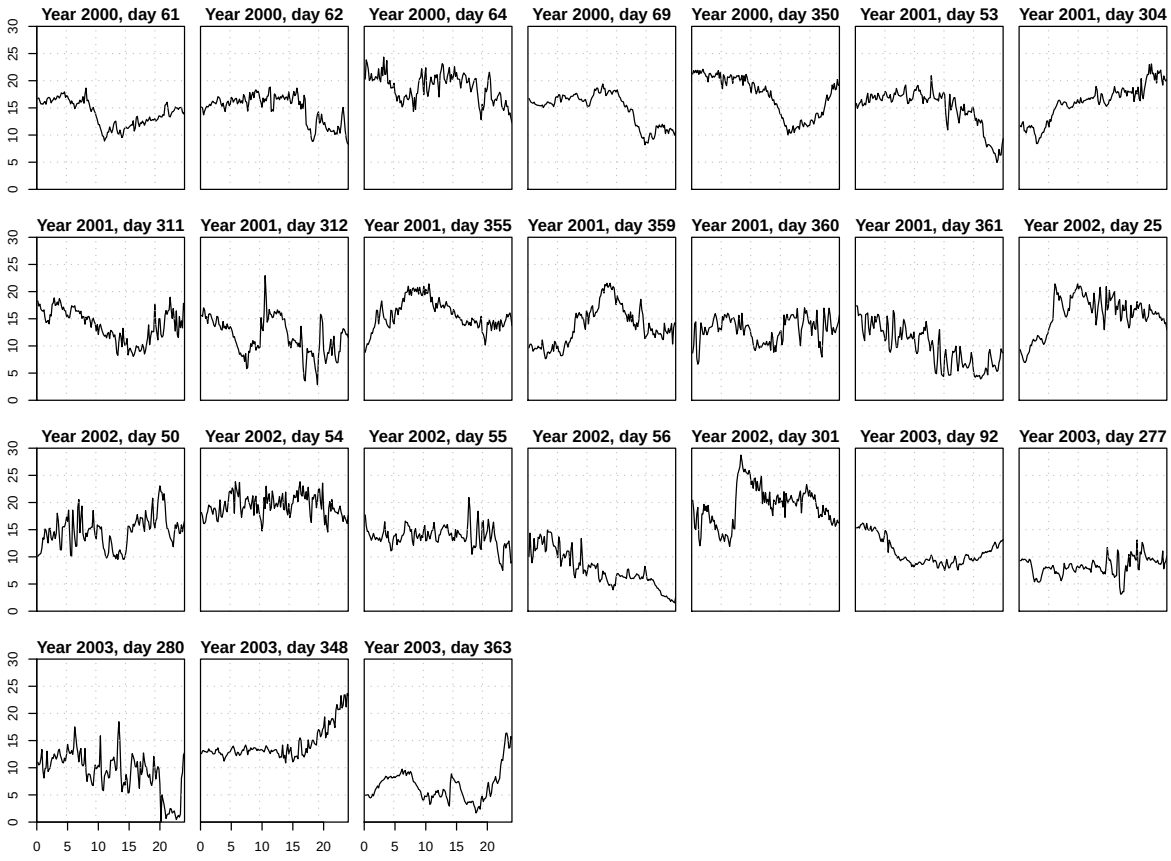


Figure 5.15: 24 hour time series of wind speed at Horns Rev M2 (centred at 00UTC) for the 24 category 15 cases that occurred during the period 2000–2003

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0	0	2	0	0	4	1	0	0	0	0	0	0	1	0	0	0	0
2	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0
4	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
5	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
6	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	2	2	2	0	0	0	2	1	0	3	0	0	0
8	0	0	1	0	0	2	1	2	1	0	0	1	0	0	0	0	0	0
9	0	0	1	0	0	1	0	0	0	2	0	0	0	0	0	0	0	0
10	0	0	1	0	0	4	1	4	0	1	0	0	1	1	0	0	0	0
11	0	2	2	0	0	5	1	2	0	2	0	1	0	0	1	0	0	0
12	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	2	0	0
13	0	0	1	0	0	8	5	6	0	2	0	1	1	3	0	4	0	0
14	0	0	2	0	0	3	3	3	1	3	0	1	0	0	0	3	0	1
15	3	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	4	0	0	0	0	1	0	1	0	0	0	2	0	0	1	0	0	0
17	0	0	0	0	0	1	1	0	1	0	0	0	0	0	2	1	0	0
18	0	0	1	0	0	3	2	0	0	1	0	1	2	0	0	1	0	0
19	0	1	0	0	0	2	0	0	1	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	2	1	1	0	2	0	0	1	2	0	3	0	0
21	0	0	2	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
22	1	4	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
23	1	1	2	0	0	0	1	0	1	1	0	2	0	0	0	0	0	0
24	0	0	2	0	0	3	5	3	1	2	0	1	1	0	1	1	1	0
25	0	0	0	0	0	3	3	2	0	2	0	1	0	1	0	0	0	0
26	0	0	2	1	0	2	0	0	0	0	0	0	0	0	0	1	0	0
27	2	1	1	1	1	1	1	1	0	0	1	0	0	0	1	0	0	0
28	0	0	0	0	0	3	0	1	1	0	0	0	1	2	0	3	1	0
29	0	0	0	0	0	2	11	6	0	1	0	4	0	1	2	1	0	1
30	0	0	0	0	0	1	0	0	0	0	0	2	3	0	1	1	0	0
31	0	0	0	0	0	5	2	1	0	0	0	0	0	0	1	0	0	0
32	0	0	0	0	0	7	0	1	0	0	0	0	1	0	0	0	0	0
33	0	2	3	0	0	0	2	0	0	0	0	0	2	0	1	1	0	0
34	0	1	3	0	0	1	0	0	1	1	0	1	0	0	1	0	0	0
35	0	0	0	0	0	2	6	0	0	0	0	4	2	1	4	1	0	0
36	0	0	0	0	0	1	2	1	0	1	0	1	2	0	0	0	1	0

Table 5.3: Comparison of the number of days during 2003 when SOMs categories (rows) and satellite classification categories (columns) occur. The satellite category columns that include OCC and the SOMs category rows relating to severe wind variability are shaded.

5.6 Open cellular convection and horizontal wind fluctuations

OCC clearly contributes to the variability of the atmosphere on length scales similar to the size of the cells. To illustrate the connection between OCC and fluctuations in the horizontal

wind speed near the surface in an explicit way, a synthetic aperture radar (SAR) image of the 10 m winds over the North Sea is shown in figure 5.16. The SAR image comes from a sensor on board the ENVISAT satellite, where 10 m wind speed is estimated based on the radar signal that is returned from the ocean signal through Bragg scattering (Christiansen et al., 2006). In this case, the image was retrieved by Risø-DTU and processed by the APL/NOAA SAR wind retrieval system developed at The Johns Hopkins University Applied Physics Lab[1].

In the top right and bottom panels of figure 5.16, the corresponding visual satellite picture and wind speed time series from Horns Rev are shown. The SAR image is valid at 0957 UTC on the 6th November 2007. The satellite picture is valid at 1015 UTC on the same day, and the time series runs from 0000 UTC to 2400 UTC. Since there was lower availability of SAR images during the main study period of the thesis (2000–2003), a case from 2007 was chosen. The meteorological mast M2 was not available at this time, so an alternative offshore mast (M7) just to the east of the Horns Rev wind farm was used instead. The wind speed is measured at a height of 70 m, and the wind direction at a height of 68 m. Although the mast is in the wake of the wind farm and will be somewhat effected by it, the large, hour scale wind fluctuations with an amplitude of up to 10 m s^{-1} are not expected to be strongly influenced by the wake.

The obvious cellular patterns in the horizontal wind speed match well with the size of the cells in the visual satellite picture. In the SAR image, the horizontal wind speed varies between about 13 and 20 m s^{-1}. The observed time series from Horns Rev shows wind speeds between 17 and 24 m s^{-1}, which are stronger than the observed SAR wind speeds because they are from a height of 70 m while the SAR wind speeds correspond to a height of 10 m.

The precise connection between horizontal and vertical wind speed fluctuations will be discussed further in chapter 10, where cross sections of flow in modelled cells will explicitly show how the surface divergence associated with the upwards vertical velocity in the cell walls and downward vertical velocity in the cell centres must be accompanied by large fluctuations in horizontal wind speed.

5.7 Discussion and Conclusions

The analysis of wind variability as a function of the synoptic flow demonstrated that there are certain large scale weather conditions that carry a high risk of severe wind variability. The SOMs analysis effectively related the large scale flow to the local conditions, and showed that there is not always a simple relationship between the local and large scale observations. For example, there are several SOMs categories with very similar mean local wind direction (see categories 4,6,11,14 and 36 in figure 5.9) which relate to quite different weather patterns. The SOMs categories that related to high risk wind variability events were generally representative of west to northwesterly synoptic flow, which the exception of category 21 where there was a low pressure centre close to the North Sea. In the case of the high-risk categories 15, 16 and 22, the synoptic scale flow has a deep northerly component, and is therefore likely to be associated with cold air of polar origin.

[1]The SAR image was provided by Merete Badger and Charlotte Bay Hasager from Risø-DTU

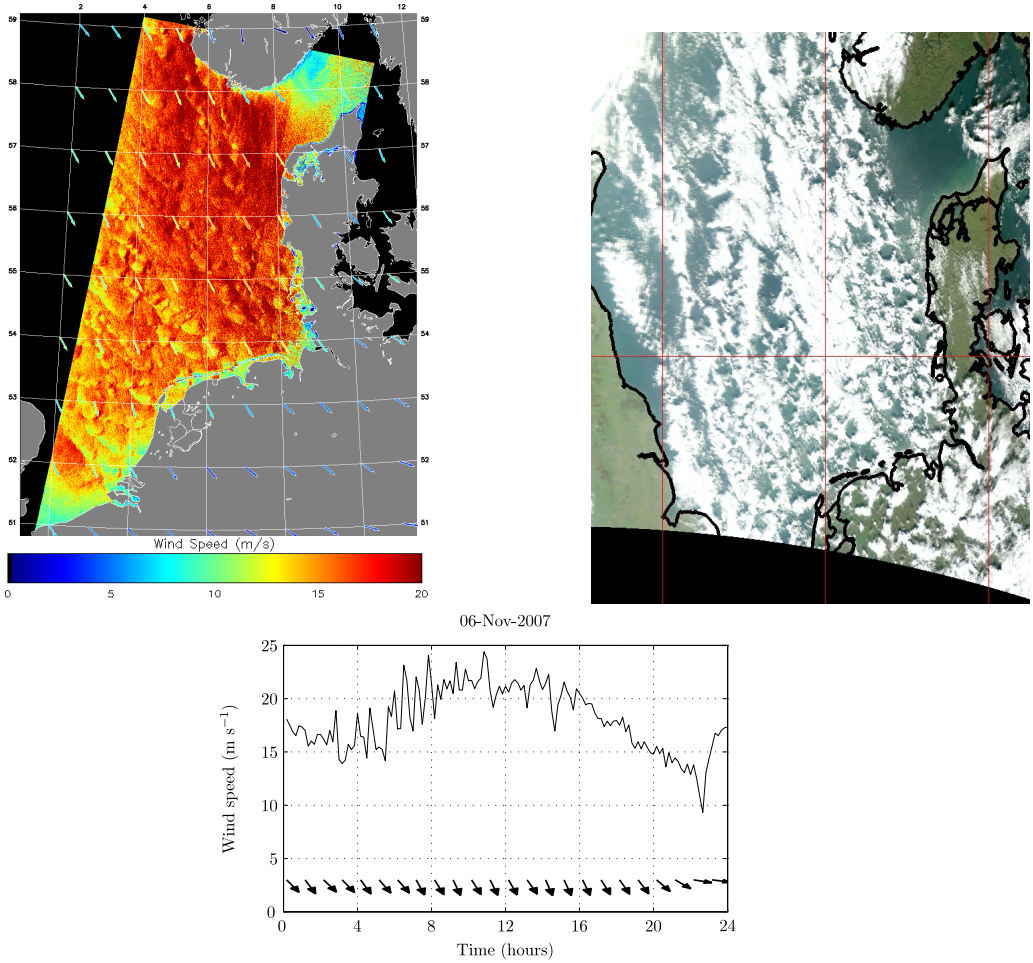


Figure 5.16: Top left: SAR image of horizontal wind speed, 0957 UTC, 6th November 2007. Top right: Visual satellite picture, 1015 UTC 6th November 2007. Bottom: Time series of wind speed at 70 m (lines) and wind direction at 68 m (arrows) from Horns Rev M7, 0000 UTC - 2400 UTC, 6th November 2007.

Like any classification system, the SOMs classification has some limitations relating to the boundaries of the categories, and to the number of categories. There will always be cases which do not fall cleanly into any category, and these cases will be placed in the best fitting category by the algorithm. No attempt to neglect these poorly fitting days was made, although it is possible that this would improve the analysis. As already discussed, there is no clear way to select the number of nodes in the SOMs array, and the final choice of a 6 $\times$ 6 array was made as a compromise between having enough resolution to distinguish between different weather types, and having a small enough array to make a meaningful classification.

Despite these shortcomings, the SOMs analysis leads to the useful observation that severe wind variability is associated with certain weather patterns, and in particular it may often be associated with cold air outbreaks. While precise local conditions are difficult to forecast using physical or statistical modelling, the synoptic scale MSLP pattern can be well modelled with even a relatively low resolution numerical weather prediction model. The analysis also demonstrates the technique of using SOMs to establish the likely large scale forcings for certain local weather outcomes, and it is probably that the technique is directly transferable to other local phenomena, some of which (such as large ramps or storms) are also of relevance

to wind energy. However, it is unlikely that all mesoscale weather phenomena show such strong links to the large scale flow.

Extensions of the SOMs analysis include using extra information as input vectors. This could include upper-air information such as 500hPa height, or more detailed surface information such as air-sea temperature difference. The inclusion of such information could help to distinguish between, for example, shallow surface features or deep features extending high into the troposphere. As a predictor, the forecast SOMs wind category is easily obtained by mapping the forecast MSLP pattern to one of the 36 categories. Since the SOMs analysis is based on large scale weather information (here, the resolution of the ERA interim reanalysis data was 1.5°), the prediction of severe wind variability using this technique does not require any high resolution or computationally expensive modelling.

The wind classes methodology also showed skill in differentiating high-risk from low-risk wind variability days. However, in terms of a potentially useful predictor of severe wind variability, this method is at risk of errors arising from small phase errors which will render the single point near Horns Rev in the wrong wind class at the single time used for the analysis. In contrast, the SOMs classification is based on the large scale weather pattern, and would generally be more robust to phase errors.

The analysis of visual satellite pictures of the North Sea region for a 365 day period in 2003 was necessarily more simplistic than the analyses based on SOMs or wind class classifications. There was a high degree of subjectivity in assigning the satellite categories, and in choosing the most representative satellite picture from several that were available each day. There was only one observation per day, since each of the Terra and Aqua satellite only pass over the North Sea region one per day, both at around noon. Furthermore, only one year of data was analysed. A more thorough analysis would use an automated method for the detection of cloud types (for example, an image recognition procedure), and would use a longer data set. The analysis was also limited by its reliance on clouds as a marker of mesoscale variability in the wind field, since it is likely that some boundary layer clouds were obscured by higher clouds, and that some mesoscale phenomena were not marked by clouds.

However, although the analysis was not statistically rigorous, some strong patterns emerged. It was found that OCC was observed on 28 days, with possible or unclear OCC on a further 29 days. Of the 19 days in 2003 which were classified as severe variability days, OCC or possible OCC covering all or part of the North Sea region was observed on 14 days. The results generally agreed with the SOMs analysis, in that OCC or possible OCC was observed on 50% of the days that occurred in the SOMs categories with the most variability days. This is much higher than the overall occurrence of OCC or possible OCC during the year, which was 16%.

Visual satellite pictures are not the only satellite observations that can be used to give information about mesoscale wind variability. For example, the ENVISAT Synthetic Aperture Radar retrieves wind speeds over the sea by relating the scattering of microwave radiation due to ocean waves to the local near-surface wind speed. Such measurements have been shown to capture small scale wind variability (Brusch et al., 2008; Christiansen and Hasager, 2006), and could be an important data set for developing climatological averages of offshore mesoscale wind variability over a large area. The example SAR image included in this chapter, in which the wind fluctuations associated with OCC are clearly visible, suggests an interesting data set for further analysis.

The analysis of satellite pictures strongly suggests that severe wind variability tends to occur when cold air flows over the warm North Sea and promotes the onset of OCC. This result is consistent with all the other results presented so far. For example, the SOMs analysis suggested preferred synoptic patterns of northerly to westerly flow driven by a low pressure centre to the north of the North Sea, which is a cold air-outbreak situation. This matches the results of the local time series analysis in chapter 4, where it was shown that severe wind variability tends to occur for wind directions in westerly to north-northeasterly sectors. Furthermore, the results are consistent with the spectrum conditioned on time of year, which showed preferred months of September to February for severe wind variability. During these months, the sea water is relatively warm compared to the cold air in a cold air outbreak, which creates the ideal conditions for the development of OCC.

OCC occurs mostly over the ocean, although Bennett et al. (2010) did recently show compelling evidence of OCC of a smaller scale over land. The results in this chapter are highly site specific, and in other locations there may be an entirely different set of phenomena that contribute to mesoscale wind fluctuations. In particular, over the land, features relating to topography (for example valley or mountain breezes) or to land-based surface heating (for example, large thunderstorms driven by diurnal heating of the land in unstable conditions) are probably more important contributors to mesoscale wind fluctuations than OCC.

Having identified OCC as an important risk factor for large amplitude mesoscale wind fluctuations, a review of the phenomenon and the related phenomena of closed cellular convection and convective rolls is presented in chapter 6.

The results of this chapter and chapter 4 suggest that there are significant and physically realistic relationships between the mesoscale wind variability and the large scale flow. This hypothesis is tested in chapter 7, where several simple statistical models are implemented. All of the models are designed to predict the mesoscale wind variability based on various parameters from a large scale weather model.

However, mesoscale meteorology is extremely complicated, and we should not expect to be able to reduce it to a simple statistical model. Therefore, in chapters 8 to 10, the possibility of modelling mesoscale wind fluctuations using a mesoscale model WRF are investigated. In particular, having identified OCC as a significant risk factor for the development of severe mesoscale wind variability, the explicit development of this phenomenon in WRF will be studied in some detail.

Chapter 6

Mesoscale shallow convection

6.1 Introduction

In chapter 5, the classification of severe wind variability according to synoptic type, satellite cloud patterns and wind classes suggested that severe wind variability can often be linked with cold air outbreaks, where cold air flowing around the rear side of a low pressure system passes over the comparatively warmer water of the North Sea. In cold air outbreaks, convective rolls and open cellular convection (OCC), which belong to the family of mesoscale shallow convection (MSC) are often observed, and are therefore obvious candidates for causing large fluctuations in wind speed. Of course, a simple correlation between these kinds of weather patterns and the occurrence of severe wind variability does not imply causality, but the hypothesis that MSC can cause large wind fluctuations also makes physical sense: MSC has the right length scale to cause variance on a time scale of about an hour, and is associated with strong surface divergence that must be associated with fluctuations in horizontal wind speed at the surface. Furthermore, it was demonstrated in chapter 5 that OCC was present during many of the occurrences of severe wind variability observed at the Horns Rev wind farm in 2003. The purpose of this chapter is to present the observational, theoretical and numerical modelling background to mesoscale cellular convection. The scientific theory discussed in this chapter underpins the work in chapters 8, 9 and 10, where the modelling of OCC is studied in some detail.

Mesoscale shallow convection was reviewed by Atkinson and Zhang (1996). Three distinct cloud forms were discussed: Open cells, which have clear centres and cloudy walls, closed cells, which have cloudy centres surrounded by areas of clear sky, and convective rolls which appear as parallel rows of clouds. They summarised a large body of research, which showed that convective rolls had been observed with wavelengths between 1 km and 13.7 km, open cells had been observed with diameters of between 2 km and 100 km, and closed cells had been observed with diameters of 1 km to 80 km. These ranges of numbers reflect a variety of observation techniques, but they do show that the three phenomena are all of the right spatial scale to potentially be able to contribute to the mesoscale variance of the wind.

Atkinson and Zhang (1996) quoted earlier work of Agee (1987), who presented a global map that suggested that OCC tends to occur as cold air flows over the warm ocean currents to the east of the major continents, while closed cellular convection tended to occur in relation to large stratocumulus fields occurring over cold water to the west of major continents. His preferred regions for OCC included an area covering the western coast of northern Europe and the North Sea, where the warm gulf stream current crosses the Atlantic Ocean towards western Europe.

Brown (1980) presented a review of convective rolls in the planetary boundary layer, and argued that convective and dynamic instabilities can both lead to such secondary flows. He suggested that in free convective situations, cells will develop, while longitudinal rolls will be produced in the presence of a moderate mean wind. Many later publications suggest that there are in fact a range of conditions when both phenomena can occur, and cells have been observed at wind speeds up to 25 m s^{-1}.

Mesoscale shallow convection over the North Sea has been examined in several studies. In 1981, observations of convective rolls and OCC were collected over the North Sea as part of the **Con**vection and **Tur**bulence experiment *KonTur*. During the experimental campaign, measurements from several aircraft flights through OCC and convective rolls were coupled with extensive surface measurements. As part of the summary of the KonTur experiment, Busack et al. (1985) analysed the frequency of open and closed cellular convection in the Norwegian Sea, just north of the North Sea, based on NOAA 6 satellite pictures. They found that OCC was observed on 78 days in 1981, in contrast to closed cellular convection which was observed on just 5 days. The cases of OCC were concentrated around late autumn, winter and early spring, with very few cases reported in the summer months. A similar result is presented in the foreword to a special publication about the experiment, where the number of OCC, closed cellular convection and convective rolls are determined for the German Bight for the period October 1976 to September 1980. During the period, there were 10 occurrences of convective rolls, 74 cases of OCC and 26 cases of closed cellular convection (Hinzpeter, 1985). These results suggest that OCC might be the dominant form of organised convection over the North Sea region.

6.2 Observational studies

Observational analysis of mesoscale cellular and longitudinal convection has been conducted in a number of studies, where authors have attempted to describe and classify the phenomena according to parameters such as cell diameter, aspect ratio and the wind and temperature gradients of the environments in which they form.

Krueger and Fritz (1961) described some of the earliest satellite observations of OCC, where pictures from the Tiros I satellite show three incidences of open cells over the Atlantic Ocean near Bermuda, the central Pacific Ocean, and near Hawaii. They reported cells with horizontal diameters of 30 – 50 miles (50 – 80 km), with clear centres surrounded by a ring or U-shape pattern of clouds. It was pointed out that such scales are difficult to detect from the ground, and therefore the phenomena had been poorly observed or neglected altogether prior to the availability of satellite images.

Agee et al. (1973) compared the observed atmospheric aspect ratios of OCC of 20:1 to those of 3:1 obtained in laboratory experiments with fluid between differentially heated plates. Statistics from satellite and ground based observations of open and closed cellular convection were given in Agee and Dowell (1974), where the average diameter of open and closed cellular convection were found to be 30 and 32 km respectively. It was reported that the diameter of the cells could be related to the height of the convection via a relation of the form $D = Ah + B$, where D is the cell diameter, h is the convection height, and A and B are constants.

Moyer and Young (1994) made a detailed study of closed cellular convection using aircraft measurements that transected 45 individual cells in the north Pacific Ocean. The 'cell-scale' flow was found by fitting a linear least squares polynomial to each cell, and the microscale contribution was considered as the deviation from the fit. They created a composite picture of the cell flow, with moist air rising in the cell centre and warm, dry air descending at the cell walls.

Several other authors have collected statistics about open and closed cellular convection, including Brümmer et al. (1986), Brümmer (1999), Bakan and Schwarz (1992) and Miura (1986). Miura (1986) used satellite pictures and atmospheric soundings over eastern Asia for a 7 year period, and presented remarkably well defined relationships between cell height and cell diameter and between aspect ratio and Rayleigh number. Bakan and Schwarz (1992) used a more extensive data set of satellite pictures over the north eastern Atlantic Ocean, and found a very large scatter in the relationship between cell diameter and cell height, with the results of Miura (1986) occupying only a narrow subspace of his results. More recently, Bennett et al. (2010) reported a case of OCC over land, although the cell diameter was much smaller than the cases observed over water.

In a case study framework, Brusch et al. (2008) presented a case study analysis of a cold air outbreak over the North Sea, and showed wind speeds derived from QuickSCAT and ENVISAT ASAR, cloud images derived from Meteosat Second Generation and MERIS FR satellites, as well as in-situ measurements from the Fino–1 platform in the North Sea, upper-air measurements from the North Sea platform Ekofisk, numerical weather prediction forecasts for the event and a surface weather chart. Rather than calculating climatological averages of the wind speed, which smoothes out mesoscale fluctuations, Brusch et al. (2008) examined the observations for the existence of mesoscale fluctuations and sought consistency between the various data sources. Visual satellite pictures showed OCC associated with the cold air outbreak, and wind fields superimposed on cloud pictures suggested that the wind speed varied regularly with the cell walls. Observed wind fluctuations had an amplitude of up to about 9 m s^{-1}, and a spectrum of the observed wind speed time series revealed a spectral peak at a period of about 1 hour.

Sikora et al. (2010) also identified OCC in the arc-like structures in SAR-derived wind speed observations, and presented average statistics for lapse rates, heat fluxes and wind shear based on 8 years of data from the northeast Pacific Ocean. They found that the phenomenon was most often observed in the colder months of the year, between September and April. They also argued that the asymmetric appearance of the cells in corresponding satellite pictures related to an arc of larger cumulus clouds around the downwind edge of the cell (corresponding to the largest wind speed, or 'squall'), and a ring of smaller cumulus clouds around the upwind edge.

6.3 Theoretical framework

Many authors have pointed out the obvious link between the laboratory case of Rayleigh-Bénard convection and OCC in the atmosphere (Agee et al., 1973; Atkinson and Zhang, 1996; Helfand and Kalnay, 1983; Brown, 1980; Cieszelski, 1998; Brümmer, 1999). Rayleigh-Bénard convection is characterised by the cellular patterns that develop when a fluid is placed between two differentially heated plates, which is a situation that is sometimes seen as analogous to air in the boundary layer between warm sea water and a capping inversion, possibly also with cloud-top radiational cooling. Rayleigh-Bénard convection is occasionally confused with Bénard-Marangoni convection which also relies on the surface tension effects at the top of an unbounded fluid - this case clearly can not be associated with atmospheric convection. The buoyancy that occurs in Rayleigh-Bénard convection as a result of the differentially heated plates is damped by kinematic viscosity and thermal dissipation. The ratio between the buoyancy and the damping is called the Rayleigh number,

$$Ra = \frac{g}{\bar{T}} \frac{(\Delta T) z^3}{\nu K}, \tag{6.1}$$

where $\bar{T}$ is the mean temperature over the convection region, ΔT is the temperature difference between the top and bottom of the fluid, z is the height of the convection, ν is the kinematic viscosity and K is the heat-conduction coefficient (Zdunkowski and Bott, 2003). Depending on the conditions of the experiment, a critical Rayleigh number can be derived, above which convection is expected to develop.

In the atmosphere, it is the eddy viscosities for heat and momentum that damp the buoyancy, instead of the kinematic viscosity and heat-conduction coefficient, because energy is removed from the fluid through turbulence, rather than directly by molecular processes. Brümmer (1999) used these ideas to estimate the atmospheric Rayleigh number for observed cases of OCC and convective rolls using estimates of the eddy viscosity for heat and momentum based on Louis (1979).

$$Ra = \frac{g}{\bar{\theta}} \frac{(\Delta\theta_{as} + \Delta\theta_{bl}) h^3}{K_M K_H}, \tag{6.2}$$

where $\Delta\theta_{as}$ is the air-sea potential temperature difference, $\Delta\theta_{bl}$ is the potential temperature difference between the top of the boundary layer and 90 m, and $\mathrm{K_M}$ and $\mathrm{K_H}$ are the estimated eddy viscosities for heat and momentum. He plotted the roll-scale and cell-scale kinetic energy, integrated over the height of the boundary layer, as a function of the Rayleigh number, and found that the kinetic energy increases with Rayleigh number according to a power law.

As discussed by (Agee et al., 1973), a major difference between the classical theory of Rayleigh-Bénard convection and mesoscale cellular convection is in the aspect ratios of the cells, which is around 3:1 for the Rayleigh-Bénard case, and typically around 20:1 – 30:1 for the atmospheric case. A suggested reason for the discrepancy is that the Rayleigh-Bénard case is governed by molecular processes, while the atmospheric cases are governed by eddy diffusivity. These ideas were further explored in (Agee and Chen, 1973), where the variation

of eddy viscosity with height was included in the solution of the Bousinesq equations. It was shown that the direction of the gradient of eddy viscosity with height could control whether open or closed cells developed.

In contrast to the theory of (Agee and Chen, 1973), Helfand and Kalnay (1983) used results from a two dimensional simulation of the Bousinesq equations to suggest that open cells occurred when convection was driven from heating in the lower part of the fluid, and that closed cells occurred when convection was driven from cooling in the upper part of the fluid. These two cases could be related to cold air passing over warmer water, or to radiative cooling from the top of a cloud layer, respectively. Krishnamurti (1975) suggested that the occurrence of open cells, closed cells or convective rolls could be related to whether the large scale motion was sinking, rising or stationary, while Ray (1986) argued that the aspect ratio of cells, and whether they are closed or open, is related to horizontal gradients in eddy diffusivity.

Several authors have tried to understand the energetics of MSC by calculating the roll-scale or cell-scale kinetic energy budget. LeMone (1973) derived roll-scale kinetic energy equations by filtering the synoptic scale and turbulent scale flow from the momentum equations, and multiplying by the roll-scale velocity. In her budget equation, the rate of change of roll scale kinetic energy, integrated over the depth of the roll, was balanced by contributions from buoyancy, shear, the Coriolis force, and the transfer of energy with the turbulent scales. She analysed data from a 444 m television tower and several neighbouring 300 m high measurement towers to produce cross sections of roll flow, and showed that shear and buoyancy are the main contributors to roll energy, while Coriolis effects and energy exchange with turbulence played a lesser roll. LeMone (1976) expanded this work by analysing the effect of roll vortices on the distribution of turbulence. Although these two works used data measured over land, the results are expected to be generally applicable over the water. Brümmer (1985) calculated a similar energy budget for convective rolls that were observed over the North Sea during the KonTur experiment, and found that of his three cases, two were shear driven and one was buoyancy-driven.

Cieszelski (1998) reanalysed aeroplane observations collected during the 1981 KonTur experiment to calculate mean flow patterns within open cells. While the observations only covered three vertical levels at about 90 m, 520 m and 850 m, the flow patterns hinted at the strong updrafts within the cell walls, and a large area of downwards vertical velocity in the cell centre. He also adapted the roll scale kinetic energy budget of LeMone (1973) to find a cell scale kinetic energy budget, and used measurements from KonTur to calculate the buoyancy and shear terms of the budget. He showed that buoyancy production was large compared to shear production.

6.4 Numerical Modelling studies

As discussed by de Roode et al. (2004), large eddy simulations (LES) of MSC are complicated by the dual requirements of a small horizontal grid spacing, and a large enough domain such that the growth of cells is not affected by the domain size itself. In other words, the domain should be able to contain a representative sample of the phenomenon. Müller and Chlond (1996) creatively responded to this challenge by using a domain that increased in size and decreased in horizontal resolution as the cells grew. As the domain increased in size,

variables from the previous, smaller domain were tiled to provide boundary conditions. They successfully simulated OCC, where the length scale of all variables, including the vertical velocity, was dominated by the cells themselves. As the possibility for larger scale computing efforts emerged, several studies simulated MSC directly on an LES grid. For example, de Roode et al. (2004) simulated a convective boundary layer, stratocumulus topped boundary layer and smoke topped boundary layers using a domain size of 25.6 × 25.6 km and a horizontal grid spacing of 100 m. The showed that for the convective boundary layer, the vertical velocity and virtual potential temperature were dominated by scales close to the boundary layer height, while the potential temperature and specific humidity were dominated by mesoscale fluctuations, and that this could occur because the mesoscale fluctuations in potential temperature and specific humidity were anti-correlated with one another such that the fluctuations cancelled when they were linearly combined to calculate the virtual potential temperature. For the stratocumulus topped boundary layer, the length scale of the vertical velocity was larger, but remained constant with time, while the length scales of all other variables grew and developed mesoscale fluctuations. Schröter et al. (2005) used a parallelized LES model to simulate MSC with a domain size of 70.4 × 70.4 km and a horizontal grid spacing of 100 m. They showed that while the temperature and liquid water path were dominated by length scales of 9 to 11 times the boundary layer height, the vertical velocity spectrum showed the most energy for length scales at 1 to 3 times the boundary layer height. This agrees with the observational study of Moyer and Young (1994), who found microscale eddies dominated over cell scale eddies in the production of vertical velocity. Schröter et al. (2005) contrasted their results with those of Müller and Chlond (1996), where the vertical velocity length scale was dominated by cell size. Furthermore, they showed that large aspect ratios of the cells only developed when latent heat release within clouds and cloud-top radiation were present in the simulation. They argue that their results show that condensation and long-wave radiation processes are therefore responsible for the existence and broadening of mesoscale convection cells.

Wang and Feingold (2009a,b) used the Advanced Weather Research and Forecasting model (ARW) with a horizontal grid spacing of 300 m to model pockets of OCC within marine stratocumulus. As they discuss, pockets of OCC within marine stratocumulus are promoted and sustained through precipitation and evaporative cooling. This is a different mechanism to the formation of OCC when cold air blows over comparatively warmer water, such as in the cases discussed in the observational studies above, but demonstrates the possibility of using WRF for modelling MSC.

6.5 Summary

There are three related phenomena that belong to the family of MSC: Convective rolls, OCC and closed cellular convection. All of them have length scales that mean that they can contribute to mesoscale wind variability, and all are observed over extensive parts of the world's oceans. A variety of theories have been proposed to explain the occurrence of open and closed cells, although there are clearly many complex processes in the atmosphere that differentiate atmospheric cellular convection from theoretical cases. Over the North Sea, several studies, together with the results presented in chapter 5, suggest that OCC is the most commonly observed form of MSC.

Chapter 7

Statistical models of wind variability

7.1 Introduction

The results of chapters 4 and 5 showed that there are strong relationships between the locally observed wind variability and the local and large scale meteorological conditions. While these results were consistent with physical arguments relating to cold air outbreaks and cellular convection, they also suggest that it might be possible to predict mesoscale wind variability using fairly simple statistical models that model the locally observed wind variability based on the large scale weather conditions. This has obvious practical outcomes, since if satisfactory models can be built based only on the large scale weather, then useful forecasts of the onset of wind variability can be made without resorting to computationally demanding high resolution numerical modelling. In this chapter, nine statistical models are developed that predict the mesoscale wind variability based on data from the ECMWF ERA Interim reanalysis (Simmons et al., 2006/2007) and on past values of the mesoscale wind variability itself.

In the approach taken in this chapter, the scalar metric of wind variability is predicted, but not the wind speed itself. The scalar metric of variability says nothing about the mean value of the wind speed, but instead shows the predicted amplitude of wind fluctuations within a certain range of frequencies. Here, the variability metric developed in chapter 4 for periods between 1 and 3 hours was chosen for analysis, but depending on the needs of system operators and other end users of wind and power forecasts, different time scales could easily be chosen. For example, knowing that the amplitude of the wind speed variations of a certain frequency are likely to increase, transmission system operators could ensure that the necessary reserve power to react on the relevant time scales was available. Many interesting models for the wind speed itself already exist, and any model of wind variability would naturally be a supplement to a model that forecasts the mean wind. These ideas have already been explored by Davey et al. (2010), who created a scalar metric of variability based on the moving average of the band-pass filtered wind speed, and then modelled it using a random forest model with inputs from the NCEP reanalysis.

Some wind speed forecasting models are also relevant to wind variability forecasting. Several recent works have built on the empirical mode decomposition (EMD) that was used extensively in chapters 3 and 4, and its more recent version, the ensemble empirical mode decomposition (EEMD). This methodology is particularly exciting, since it presents the possibility of forecasting different modes of variability (which may well relate to different physical processes in the atmosphere) in isolation from one another, before recreating the whole time series through linear superposition. For example, Xingjie et al. (2009) decomposed a wind speed time series using the EMD, then added them in groups to create the low frequency and high frequency parts of the time series, which together with the low frequency trend could be added to recover the original time series. They forecast the low frequency oscillating part and the low frequency trend with an ARMA model, and used an Elman network model for the non-linear high frequency part. Similarly, Yu et al. (2008) used EMD to decompose time series of crude oil prices, then forecast each component using a feed forward neural network and recombined the components using an adaptive linear neutral network. Such methods have great potential for the problem of wind energy forecasting, and although they are not explored here, they are certainly interesting areas for future work. If the modes of variability in the time series are to be forecast directly, it seems essential that the EEMD rather than the EMD is used, since as discussed in Wu and Huang (2009), the problem of mode mixing in the EMD means that different frequencies can appear within the same mode.

Also of relevance to the problems of forecasting the fluctuating wind speed is the work of Pinson and Madsen (2010), who observed that the fluctuations in wind speed for time scales of a few hours underwent sudden changes in characteristics at the Horns Rev wind farm. They designed and implemented Markov switching auto-regressive (MSAR) models, where the *regime-state* of the wind was an unobserved Markov process, and coefficients were estimated adaptively. It seems probable that their regime-states could be related to the conditions for the occurrence of severe wind variability that were defined in chapter 4.

One of the challenges of statistical modelling of meteorological variables is the treatment of the circular wind direction variable. It could be treated as a categorical variable, or used as a conditional variable for direction-dependent coefficient estimation, but these strategies ignore the fact that it is actually a continuous variable with a strong auto-correlation structure. Hering and Genton (2010) suggest using the sine and cosine of the wind direction as a predictor. In the models presented here, the circular nature of wind direction is ignored, and it is treated in the same way as the other continuous variables, although reformulating the model using the sine and cosine of direction suggest that a small improvement may be gained in this way.

The predictors for the statistical models of wind variability in this chapter are based on results from previous chapters, and are: 10 metre wind speed, 10 metre wind direction, geostrophic wind speed, geostrophic wind direction, mean sea level pressure, air-sea temperature difference, SOMs category and pressure tendency and the previous values of the variability metric itself. All of the predictors are either available directly or can be derived from the ERA Interim reanalysis. The pressure tendency was calculated as the difference between consecutive analyses separated by 6 hours, and the geostrophic wind speed and direction were calculated from the gradient of the mean sea level pressure field.

7.2 Real-time calculation of observed variability

In chapter 4, time series of wind variability for periods of 1–3 hours were calculated. Calculation of the wind variability was based on wind speed time series of length 1 year, so that any end effects the arose through the empirical mode decomposition and Hilbert transform were not significant.

The focus of this chapter is in forecasting the wind variability. For time series models that use past values of the time series itself, the end effects of the calculation of variability may be very important. If the model prediction of the variability at time step t_k, $\widehat{V}(t_k)$, uses the observed value of the variability at a previous time step time t_{k-n}, $V(t_{k-n})$, then it must be possible to calculate the variability at time step t_{k-n} without using the observed wind speed that occurs in the future after time t_k. That is, it is necessary that a point close to the end of the variability time series is reasonably accurate and not strongly influenced by the end effects of the empirical mode decomposition and the Hilbert Transform.

The statistical models in this chapter are built using the ECMWF Interim reanalysis data, which is available at a temporal resolution of 6 hours, while the variability time series calculated in chapter 4 is has a temporal resolution of 10 minutes. The variability time series were therefore averaged to 6 hour time resolution using centered averaging. This means that the average variability at time step t_k is an average of the local variability between times $t_k - 3$ hours and $t_k + 3$ hours. Similarly, the average variability at time step t_{k-1} is an average of the local variability between the times $t_{k-1} - 3$ hours and $t_{k-1} + 3$ hours, or in other words between 9 and 3 hours before time t_k.

This issue is illustrated in figure 7.1. The variability with 10 minute time resolution calculated using the whole time series is shown in black, and the variability calculated using wind speed observations only up to time t_k is shown in red – that is, the red curve is affected by end-effects, while the black curve is not. The curves diverge in the final 3 hours before time t_k, but before this they are very similar.

The averaging period for the average variability at time step t_{k-1} is indicated by the vertical dotted lines, and the sections of both the 10 minute variability with end-effects and without end-effects that are included in the average are shown in bold. The variability without the end effects is more correct, but calculating it involves knowing future values of the wind speed, making it useless for inclusion in a real-time statistical model. The variability with and without end effects are therefore named as the *real-time variability* and the *a posteriori variability* respectively.

The *a posteriori* variability and the real-time variability were calculated for each 6 hour period in January 2002, to test whether it is reasonable to build models that require an input of the past variability calculated in real-time. A scatter plot of the two values is shown in figure 7.2. The plot suggests a close relationship between the real-time variability and the *a posteriori* variability, particularly since new and more effective ways of handling the end effects of the HHT may be developed in the future.

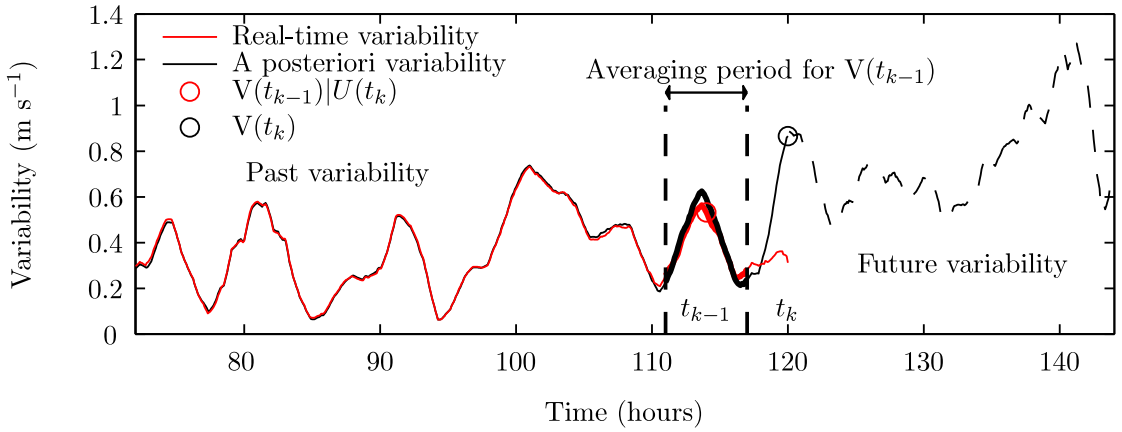


Figure 7.1: Real-time variability (red) and *a posteriori* (black) variability time series. The solid and dashed black line shows the *a posteriori* variability before and after time t_k respectively, which would be the value of the variability if there were no end effects. The red line shows the variability calculated using wind speed observations only up to time t_k. The variability metric has a time resolution of 10 minutes, which is averaged to a resolution of 6 hours to match the time step of the other inputs to the statistical model. The average variability metric at time t_k is calculated as the average of the local variability between $t_k - 3$ hours and $t_k + 3$ hours, and the average variability metric at time t_{k-1} is calculated as the average of the local variability between $t_k - 9$ hours and $t_k - 3$ hours. The red dot indicates the variability at time t_{k-1} given the availability of wind speed observations up to time t_k ($V(t_{k-1})|U(t_k)$).

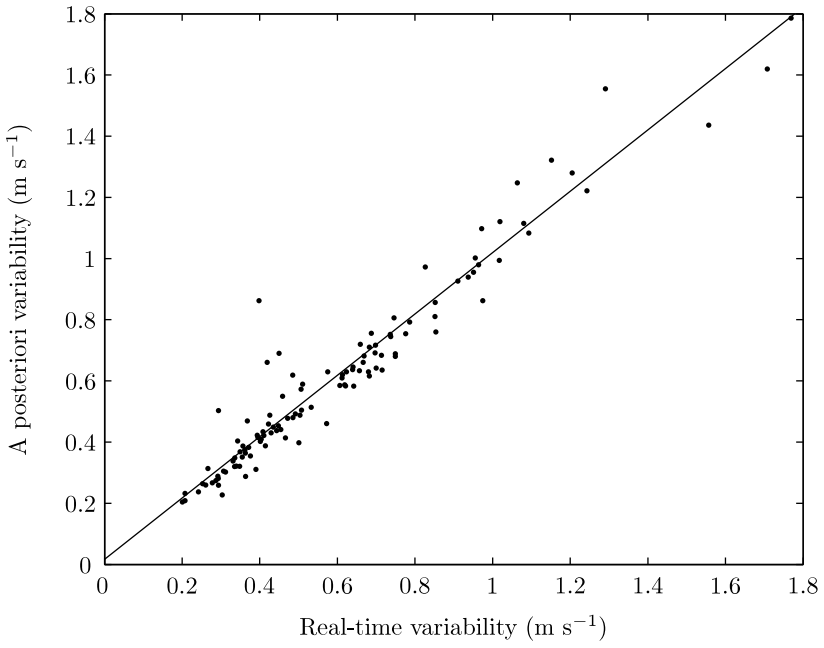


Figure 7.2: Real-time and *a posteriori* variability time series

7.3 Description of models

7.3.1 Linear models

In a linear model, the response variable (here, the variability, $\widehat{V}(t)$) is predicted as a linear combination of one or more predictors (equation 7.1).

$$\widehat{V}(t) = \alpha_0 + \sum_{n=1}^{N} \alpha_n A_n(t) \tag{7.1}$$

where α_n are the constant coefficients, A_n are the predictors, N is the number of explanatory variables included in the model, and the errors are assumed to be normally distributed with constant variance.

The simplest kind of linear model is a persistence model, where the predicted value of the time series at time $t+1$ is related to the observed value at time t with a coefficient of 1. This is equivalent to saying that the value of the time series now is the same as what it was one time step ago. In this chapter, a one-step-ahead persistence model is defined and considered as a benchmark for the more complex models.

Many variations on the basic linear model are possible. Of the configurations that were tested, three are presented here. The first model (L1) is a simple linear model using the predictors 10 metre wind speed, 10 metre wind direction, geostrophic wind speed, geostrophic wind direction, mean sea level pressure, air-sea temperature difference, SOMs category and pressure tendency. The SOMs category was included as categorical data, since the absolute numbers of the SOMs categories are arbitrary. Inclusion of the categorical SOMs data means that 35 binomial variables are introduced into the regression (one for each of 35 SOMs categories. The 36th category is taken into account by the case where the remaining 35 values are zero).

The second linear model (L2) is the same as the first, except that the previous observed value of the variability was included as a predictor,

$$\widehat{V}(t) = \alpha_0 + \beta V(t-1) + \sum_{n=1}^{N} \alpha_n A_n(t), \tag{7.2}$$

where β is the constant coefficient of the variability at the previous time step.

The third linear model (L3) was formulated as a linear relationship between the log of the variability and the same explanatory variables as L2, in an attempt to give a better model of the infrequent, large spikes in variability. This formulation also naturally enforces the restriction $\widehat{V}(t) > 0$.

$$\log\left(\widehat{V}(t)\right) = \alpha_0 + \beta V(t-1) + \sum_{n=1}^{N} \alpha_n A_n(t)$$
$$\widehat{V}(t) = e^{\left(\alpha_0 + \beta V(t-1) + \sum_{n=1}^{N} \alpha_n A_n(t)\right)}. \tag{7.3}$$

7.3.2 Generalised Linear Models

Generalised linear models (GLMs) (McCullagh and Nelder, 1989) are a generalisation of the basic linear model, where the random component of the response variable is not necessarily normally distributed, and where the expected value of the response variable ($\widehat{V}(t)$) is related to the linear combination of predictors via a link function (equation 7.4). The random component of the response variable can take important distributions such as the normal, Poisson and gamma distributions. Amongst the possible distributions and link functions that were tested, the Poisson distribution with the identity link function gave the lowest model error - that is, the errors in $\widehat{V}$ are assumed to follow a Poisson distribution (thus ensuring that the variability is positive), and the mean of the distribution is related to the linear combination of predictors via the identity function.

$$\widehat{V}(t) = g^{-1}\left(\alpha_0 + \sum_{n=1}^{N} \alpha_n A_n(t) + \beta V(t-1)\right)$$
where
$$g(\eta) = \eta \tag{7.4}$$

7.3.3 Generalised Additive Models

Generalised additive models (GAMs) (Chambers and Hastie, 1992) are a further generalisation of the GLM. Like the GLM, the GAM relates the response variable to the predictors through a non-linear link function. However, the argument of the link function is now a linear combination of non-parametric functions of each of the predictors. The non-parametric functions are fitted to each of the predictors using locally weighted scatter plot smoothing (LOESS) (Venables and Ripley, 2002), where at each point on the abscissa of the scatter plot, weighted least squares fitting is used to fit a local polynomial to a segment of the data, with the highest weights applied to the nearest data points. Again, a variety of link functions are possible. As for the GLM, the Poisson distribution with the identity link function gave the lowest RMSE.

$$\widehat{V}(t) = g^{-1}\left(A_0 + \sum_{n=1}^{N} f_n(A_n(t)) + h(V(t-1))\right)$$
where
$$g(\eta) = \eta \tag{7.5}$$

The non-parametric functions, which relate each of the predictors to the response variable are shown in figure 7.3. The SOMs category was included as a categorical variable in the same way that it was in the linear model, and the pressure tendency was included as a simple linear predictor, since its LOESS fit was very close to linear anyway. The LOESS fit functions should actually mirror the climatological results that were presented in chapter 4, with a few important differences. Firstly, the functions relate the observed variability to the meteorological variables from the large scale analysis, while the results in chapter 4 related the observed variability to the locally observed meteorological conditions. Secondly, the non-parametric fitting is not perfect, particularly where there are few data. For example, in the fit between air-sea temperature difference and variability, it appears that there is an increase in variability for very high air-sea temperature differences, but this is actually a result of very few data points having an unreasonably large impact on the fit. However, the plots in figure 7.3 are generally in agreement with the results already presented, since the highest contributions to the variability occur for west to northwesterly flow directions and for higher wind speeds. The different slopes seen in the ten metre wind speed and geostrophic wind speed suggest that the model is actually over-fitted.

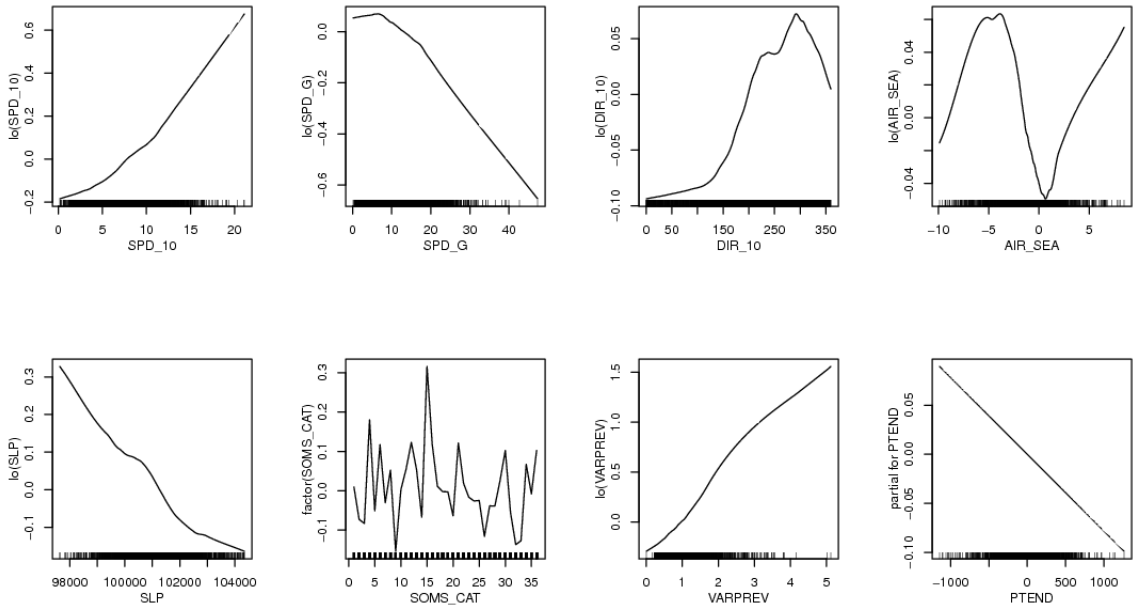


Figure 7.3: Non-parametric fits for the predictors in the GAM model. The function between pressure tendency and variability is a linear fit, and the SOMs category is treated as a categorical variable. The dashed lines along the horizontal axes show the density of data for each value.

7.3.4 ARIMA model

ARIMA models relate the response variable at time t to previous observed values of the response variable and to current and past values of the model uncertainty (Box and Jenkins, 1970). The number of lags of the modelled variable and the model uncertainty that are included in the model determine the *order* of the autoregressive and moving average components of the model respectively, and are chosen by consideration of the auto-correlation and

partial auto-correlation functions.

The ARIMA model could be formulated using a training period to calculate the constant coefficients and a testing period to apply and evaluate the model, as for the linear model, but it was found that this strategy gave extremely poor results, mainly because the model was simulating variance in the time series that looked realistic, but was often out of phase and of incorrect magnitude. Instead, the model was developed as an updating model, with the coefficients re-calculated for each time step, with a 'memory' period of three weeks.

After considering the auto-correlation function and partial auto-correlation function for a selection of three week periods, an ARIMA model of order (2,1,1) was chosen. Amongst other considerations, one can see that after one step differencing is applied, there are significant lags at one and two time steps in the sample auto-correlation functions for three different three week periods in figure 7.4, thus motivating the choice of the ARIMA model order. That is, the autoregressive component was of order 2, one step differencing was applied, and the moving average component was of order 1. The structure of the ARIMA (2,1,1) model is given in equation 7.6, where multiplication by the operator B^n denotes the value of the time series at a lag of n steps. $V(t)$ is the time series that is being modelled, ϵ are the forecast residuals and ϕ_n and θ_n are the coefficients applied to the lagged values of the time series and the lagged and current values of the forecast residuals. The one step differencing in the equation is included through the $(1 - B)$ term on the left hand side of the equation.

$$(1 + \phi_1 B + \phi_2 B^2)(1 - B)V(t) = (1 + \theta B)\epsilon(t) \tag{7.6}$$

Clearly, the values of the forecast error, $\epsilon(t)$ that have not yet happened cannot be included in predictions based on the model. Following Box and Jenkins (1970), these values are set to zero. The one, two and three step ahead predictions from the model are then formulated as:

$$\begin{aligned}
\hat{V}(t+1)|V(t) &= -\phi_2 V(t-2) - (\phi_2 + \phi_1)V(t-1) - (1 + \phi_1)V(t) + \theta\epsilon(t) \\
\hat{V}(t+2)|V(t) &= -\phi_2 V(t-1) - (\phi_2 + \phi_1)V(t) - (1 + \phi_1)\hat{V}(t+1) \\
\hat{V}(t+3)|V(t) &= -\phi_2 V(t) - (\phi_2 + \phi_1)\hat{V}(t+1) - (1 + \phi_1)\hat{V}(t+2)
\end{aligned} \tag{7.7}$$

As an example, the auto-correlation functions and partial auto-correlation functions for three different three week periods with one hour differencing applied are shown in figure 7.4.

After the ARIMA (2,1,1) model was fitted, there should be no significant lags left in the auto-correlation function and partial auto-correlation function of the residuals, which was generally the case, as demonstrated by the auto-correlation function and partial auto-correlation function for the same three week periods that are demonstrated in the previous example (figure 7.5). However, a shortcoming of this implementation of the model is that an ARIMA model of order (2,1,1) will not be equally applicable for all three week periods. For example, during some periods that have a strong diurnal cycle, it might be appropriate to include a diurnal component with a lag of 24 hours, and for periods when the synoptic cycle dominates, a seasonal component of around three days might be beneficial. There is scope for further

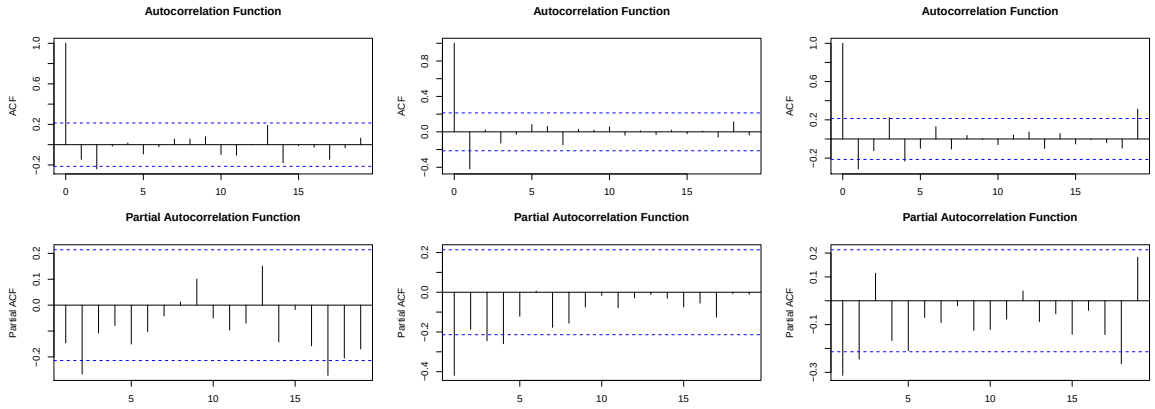


Figure 7.4: Auto-correlation functions and partial auto-correlation functions for three different three week periods during the testing period. Left to Right: three week periods in January, April and July respectively

work in developing ARIMA models for prediction of wind variability, particularly with regard to updating coefficients. However, the purpose of this work is to show the skill of some widely used statistical models, thus demonstrating the predictability of wind variability from a statistical perspective.

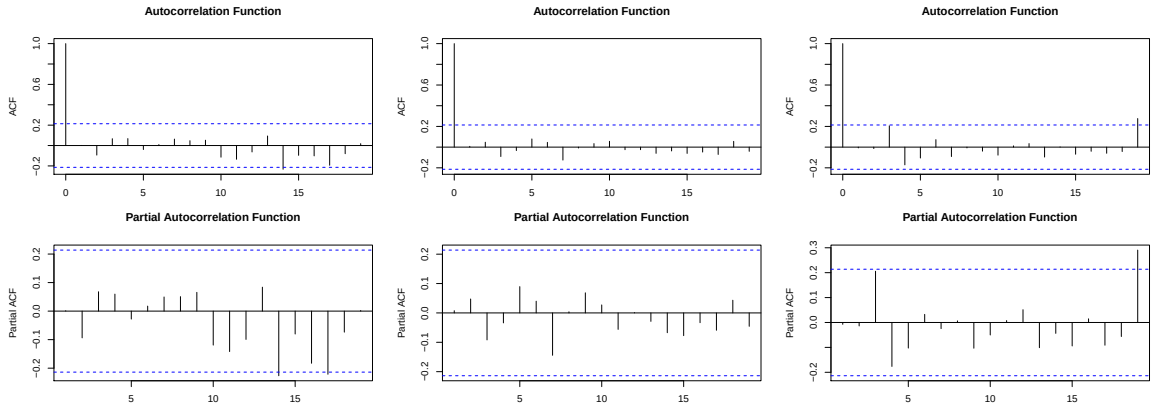


Figure 7.5: Auto-correlation functions and partial auto-correlation functions for the model residuals for the same three periods used in figure 7.4

7.4 Results and comparison of models

The first model that was tested was a one step ahead persistence model, for which the RMSE was 0.48 m s^{-1}. This is considered the benchmark against which the other models are tested.

A sample of the test time series for the three different linear models (L1, L2 and L3) is shown in figure 7.6.

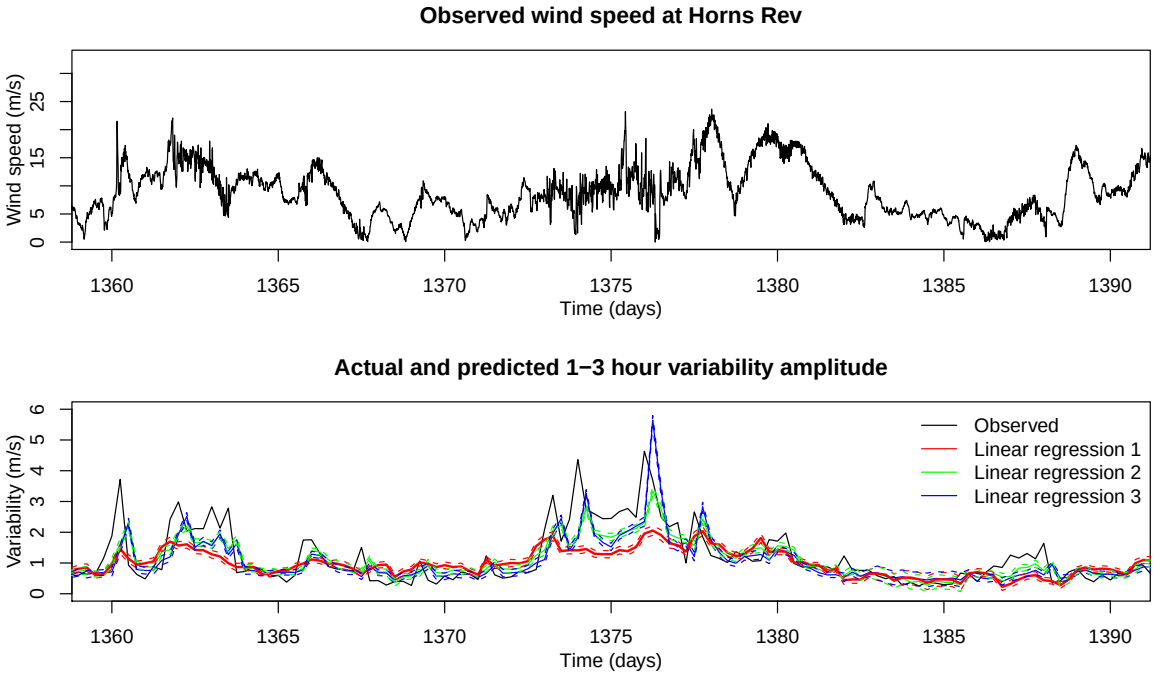


Figure 7.6: Observed wind speed at Horns Rev (top panel) and forecast and observed 1-3 hour variability time series (lower panel). Linear regression 1 uses only the reanalysis data and does not take past values of the variability time series into account. Linear regression 2 includes the previous value of the variability time series as a predictor, and for linear regression 3, the log of the variability time series is used as the response variable. The dashed lines show the point wise 95% confidence interval for each of the models.

The estimates of the coefficients for L2, which included the past value of variability as a predictor, are shown in table 7.1. The results of the t test, which is a test of whether the estimate is different from zero (accounting for the standard error and sample size), suggest that all of the predictors apart from the geostropic wind are probably significant in the model. In addition, there are the 35 extra variables arising from inclusion of the 36 SOMs categories as binary variables. For all of these variables, $Pr(>|t|) > 0.02$, suggesting that they were not very significant, with the exception of SOMs category 15, which stood out with $Pr(>|t|) = 0.0007$. SOMs category 15 was the category that was already identified in chapter 5 has being the strongest risk category for severe wind variability events, so this result is consistent with previous results.

The root mean square errors for the three models were 0.479 m s^{-1}, 0.414 m s^{-1} and 0.433 m s^{-1} respectively. This means that the best performing model was the one that included the previous value of the variability as a predictor, but was not designed using the log of the variability as the response variable.

A sample period of the GAM and GLM models are shown in figure 7.7, together with L2, the best of the three linear models. Although the RMSE for the GAM and GLM models were slightly better (0.409 and 0.413 respectively) than the linear models, the time series show that there was actually very little increase in skill as a result of the considerable increase in complexity of the GAM and GLM models.

	Estimate	Std. Error	t value	$Pr(>\|t\|)$
Intercept	9.177	1.193	7.692	1.98×10^{-14}
U_g	5.008×10^{-3}	2.945×10^{-3}	-1.701	0.0891
θ_g	8.559×10^{-5}	1.226×10^{-4}	0.698	0.485
U_{10}	2.039×10^{-2}	5.603×10^{-3}	3.639	0.000279
θ_{10}	4.281×10^{-4}	1.210×10^{-4}	3.539	0.000408
ΔT	-1.256×10^{-2}	3.326×10^{-3}	-3.776	0.000163
MSLP	-8.782×10^{-5}	1.166×10^{-5}	-7.534	6.53×10^{-14}
ΔP	-1.038×10^{-4}	3.230×10^{-5}	-3.214	0.0001323
$V(t-1)$	4.648×10^{-1}	1.659×10^{-2}	28.012	$< 2 \times 10^{-16}$

Table 7.1: Parameter estimations and significance levels for the continuous variables in the simple linear regression model

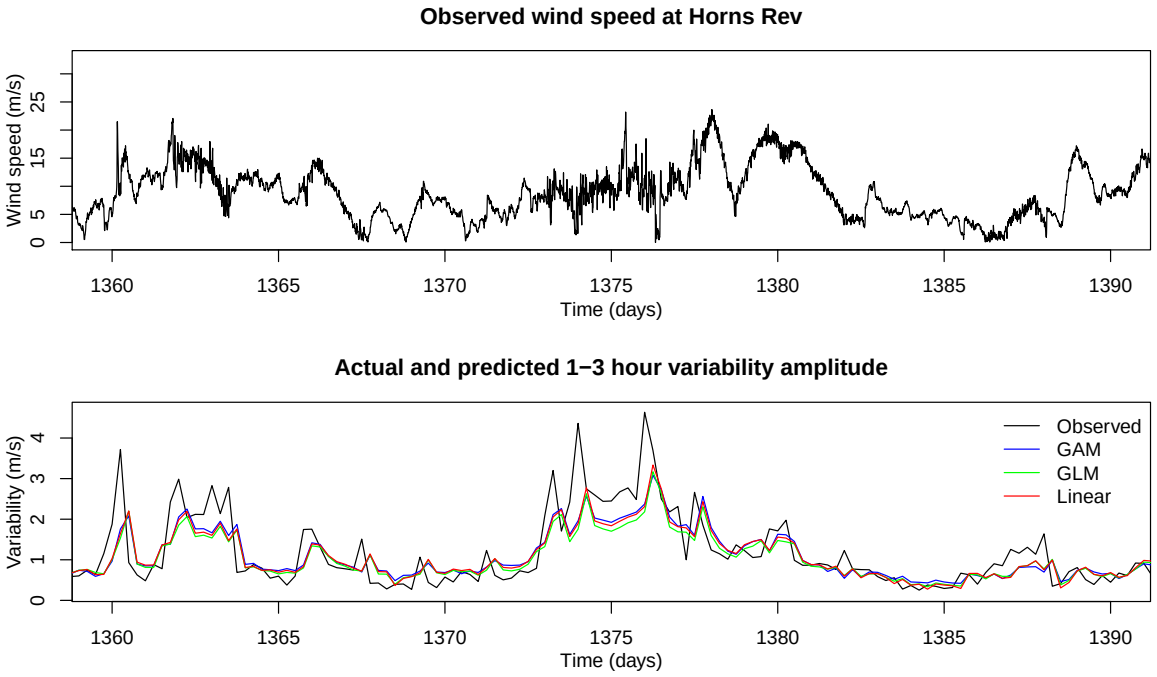


Figure 7.7: 30 day sample of the predicted and observed variability time series for the linear, GLM and GAM models (lower panel). The original wind speed time series is shown in the upper panel.

The ARIMA one, two and three step ahead models for the same sample period are shown in figure 7.8. Since the time steps of the model are 6 hours, these relate to 6, 12 and 18 hour outlook forecasts. These models have a very different character to the predictions from the linear, GAM and GLM models, since they are not relying on the smooth fields of the large scale weather model to develop variability. The three models clearly have a lag which increases with the forecast look-ahead time, due to the autoregressive component in the model which causes it to follow the previous observed states. As a result of this increasing lag, the models are very similar to one another, but offset by 6 hours. However, in terms of amplitude, the ARIMA models are better than the linear, GAM and GLM models, but the correct amplitude is only a lagged reaction to the observed values. The RMSE of the one,

two and three step ahead models were 0.59, 0.58 and 0.62 respectively, which shows that they are actually considerably worse than the persistence model.

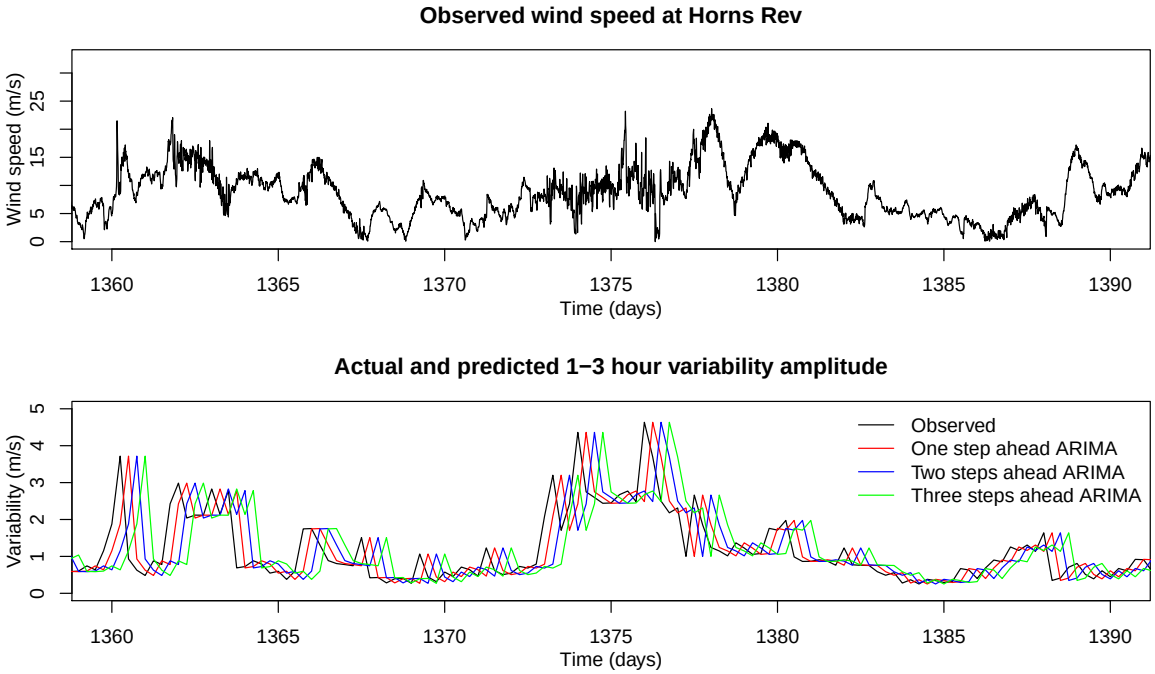


Figure 7.8: 30 day sample of the predicted and observed variability time series for the ARIMA one, two and three step ahead forecasts (lower panel). The original wind speed time series is shown in the upper panel.

Scatter plots for the linear, GLM, GAM and ARIMA one and two step-ahead models that were considered in this chapter are shown in figure 7.9. The model with the lowest RMSE is the GAM model, although the difference between the linear model with an auto-correlation term, and the GLM and GAM models is not very large. All models struggled to produce extreme values of variability. It is very likely that the RMSE could be reduced still further with the application of more sophisticated models. However, the goal of this chapter is not to develop the best possible statistical model, but to show that there is potential skill in predicting mesoscale variability based only on the large scale atmospheric conditions.

7.5 Discussion and Conclusions

In this chapter, several standard statistical models were implemented for the prediction of wind variability based on its own autocorrelation and on a set of external predictors from the ERA Interim reanalysis. The predictors included basic variables like wind speed at a single grid point that were extracted directly from the analyses, as well as more complicated derived variables such as the SOMs category and the geostrophic wind speed and direction. Inclusion of the SOMs category meant that the models could incorporate the effects of the large scale weather patterns. The models showed limited skill, with the best model giving only a 15% improvement on the persistence forecast.

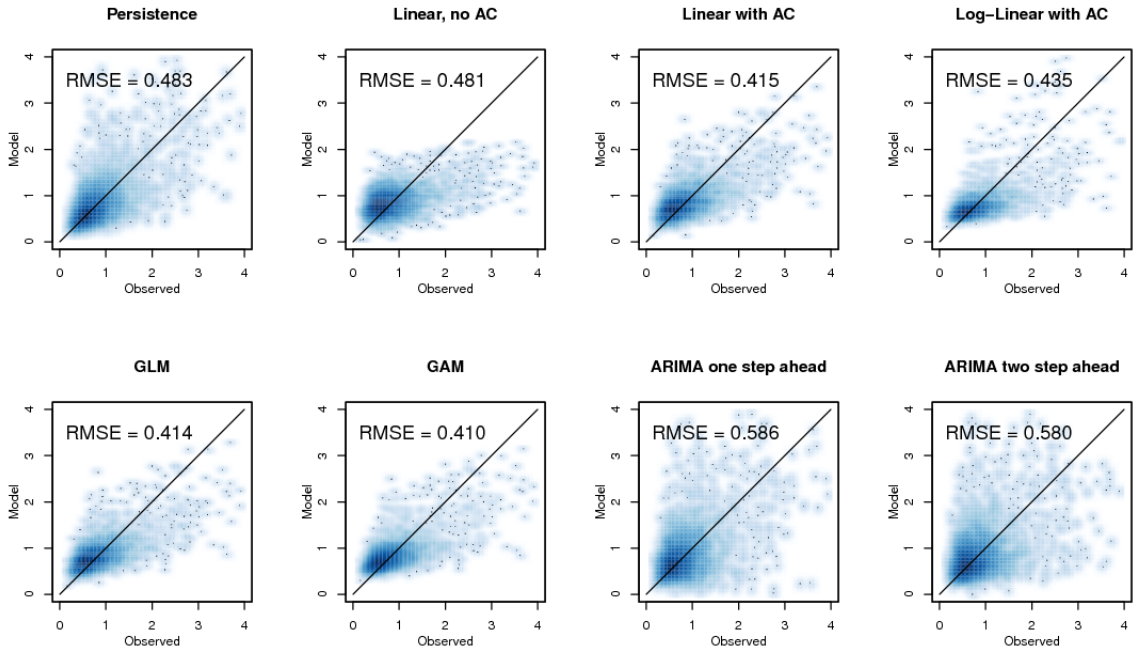


Figure 7.9: Scatter plots of modelled against observed wind variability for the eight models considered here. All models are trained on two year period 2000–2001 and verified on the two year period 2002–2003, expect for the persistence model which did not receive any training.

The inputs to the statistical models came from the ERA Interim 1.5° resolution analysis, and this only demonstrates one aspect of the predictability of mesoscale variability based on the large scale flow. To predict mesoscale variability in a realistic setting, *forecasts*, not *analyses* are required, and these forecasts will also have errors in them. However, the skill of large scale weather models is a separate problem to the predictability of mesoscale wind variability from the large scale weather conditions, and this is not addressed here.

The ERA Interim reanalysis is designed to provide a consistent time series of analyses, which means that the same model and assimilation system was used for the whole data set. However, the observational data that is used in the assimilation system necessarily changes with time, amongst other things due to the changing availability of satellite data, and this calls into question the validity of calculating a set of constant coefficients for a training period and applying them to a testing period. For example, it was visually obvious that the statistical properties of the sea surface temperature data underwent a change during the 4 year period. These issues point to the necessity of adaptively re-estimating the model coefficients using a sliding window.

It is probable that more sophisticated statistical models could bring about small improvements in the predictability of mesoscale wind variability based on the large scale weather patterns. For example, self exciting threshold autoregressive models (SETAR), which use conditional parameter estimation based on values of the time series itself, smooth transition autoregressive models (STAR), where the parameters change smoothly between different states according the time series itself or some external conditions, or markov switching autoregressive models (MSAR), where the parameters changed between different states according to the unobserved regime of the wind, could yield improvements to the simple ARIMA models implemented in this chapter, particularly considering that the characteristics of the

variability have already been shown to change with season and synoptic pattern. Furthermore, transfer function models that take the cross-correlation structures of the various input variables into account could also improve the results.

However, the modest improvements that were obtained by increasing the complexity of the models to from simple linear regression models to the GAM and GLM models suggest that even with very sophisticated statistical modelling, the prediction of mesoscale variability based on the large scale flow will not reach a high level of skill. This is because the mesoscale processes that cause the variability simply do not exist in the large scale model, and the various controls on their onset and decay are sensitive to small changes in the atmospheric forcing. Certain aspects of the mesoscale wind variability can probably only be modelled using a physical model that downscales the large scale weather model to a small domain, taking into account realistic boundary layer physics, microphysical processes, and higher resolution spectral information. These ideas are explored in chapters 8 to 10, where the predictability of mesoscale wind variability using the mesoscale model WRF is explored. In particular, the chapters focus on the phenomenon of OCC, since this has already been identified as a strong risk factor for the development for large, hour-scale wind fluctuations.

Chapter 8

Forecasting mesoscale wind variability with a mesoscale model

8.1 Introduction

Numerical weather prediction (NWP) models simulate physical processes in the atmosphere by integrating a discretised version of the Navier Stokes equation, together with expressions for conservation of mass, heat and moisture, forward in time, based on knowledge of the initial and boundary conditions of the model domain. Depending on the model's horizontal grid spacing, certain physical processes in the atmosphere cannot be resolved explicitly, and must be parameterised. For example, for a horizontal grid spacing larger than about 5 km, effects of cumulus convection must be parameterised, while for smaller horizontal grid spacing, the effect of cumulus convection is at least partly explicitly resolved (Molinari and Dudek, 1992). Similarly, since turbulence is not resolved in any NWP models, the effects of turbulent mixing in the vertical direction must be parameterised through the planetary boundary layer (PBL) scheme.

Spatial scales that can give rise to hour-scale wind fluctuations are in the order of tens of km, so it seems likely that hour scale fluctuations can be captured with a high resolution mesoscale model. For example, the cells that were seen in the satellite analysis in chapter 5 had diameters of around 20–50 km. If, as discussed in Skamarock (2004), attenuation in a mesoscale model starts for spatial scales smaller than about $7\Delta x$, then these structures should be just captured by a model with horizontal grid spacing of 2 km.

For the application of short term prediction for wind energy, it is very important to understand the skill of the model in predicting mesoscale wind variability. NWP models are generally used as input to short term wind and power prediction models such as WPPT (Nielsen, 1999) and Prediktor (Landberg, 1999, 2001), and variability in wind speed can lead to large fluctuations in predicted power, particularly if they fall within the steepest part of the wind farm power curve. As higher resolution NWP models become feasible for operational use by end users such as wind power forecasters, it becomes necessary to understand whether the variability in the model reflects a physically realistic stochastic outcome of the

predicted weather. Furthermore, simply applying standard verification methodology, such as the root mean square error, to forecasts from high resolution models will mask any real skill in the model, since forecasts with realistic mesoscale variability but small phase errors will be penalised heavily compared to smoother forecasts that do not contain any mesoscale variability at all (Mass et al., 2002).

In this chapter, the ability of WRF to reproduce mesoscale wind fluctuations is tested for a set of case studies. Although the real skill of the model can only be assessed through statistical analysis of long time series, it is also useful to study particular cases in some detail. Modelled fields of vertical velocity are compared with satellite pictures, and time series observations at Horns Rev are used to verify the model in the temporal domain. A new online metric of wind variability is added into the WRF model to calculate the variance at the end of each simulation hour.

8.2 The Weather Research and Forecasting model

The Weather Research and Forecasting model (Skamarock et al., 2008) is a state of the art modelling system that can be run at a large range of scales. The whole WRF modelling system includes a dynamical solver, a physics package, variational data assimilation module and an atmospheric chemistry module, and can be used for operational or research purposes. Here, only the dynamical solver and the physics package are used with model version 3.2 and the ARW core. Boundary conditions and initial conditions are downloaded from the archive of GFS-FNL analyses (dataset ds083.2 published by the CISL Data Support Section at the National Center for Atmospheric Research, Boulder, CO, available online at http://dss.ucar.edu/datasets/ds083.2/). Boundary conditions are updated every six hours, and no data assimilation is adopted. The WRF system is modular, allowing users to choose different parameterisations and physical and dynamical schemes through run-time namelist options. The systematic and accessible code makes it relatively simple for users to make minor changes or implement new features.

8.3 Cases

Eight cases of large amplitude mesoscale wind variability were chosen for simulation in WRF. The cases were all chosen as days where the average daily 1-3 hour period variability was in the top 5th percentile, the same definition of extreme variability days that was used in chapter 5. Cases were chosen qualitatively from the 73 extreme days during 2000–2003, and were chosen to discard cases where a single major wind change dominated the time series. In addition, a ninth case was chosen as a control, which was a day in the bottom 2nd percentile. Time series observations of wind speed at a height of 62 m and wind direction at a height of 43 m from Horns Rev I for each of the cases are presented in figure 8.1, where it is seen that the wind direction was from the west to northwest for all eight cases. Satellite picture for the cases are shown in figure 8.2, where OCC is evident in all of the cases, except for the last case which was the control case of very low wind variability. However, the cases were chosen because of their large hour-scale wind fluctuations, and the fact that all had OCC was not intentional.

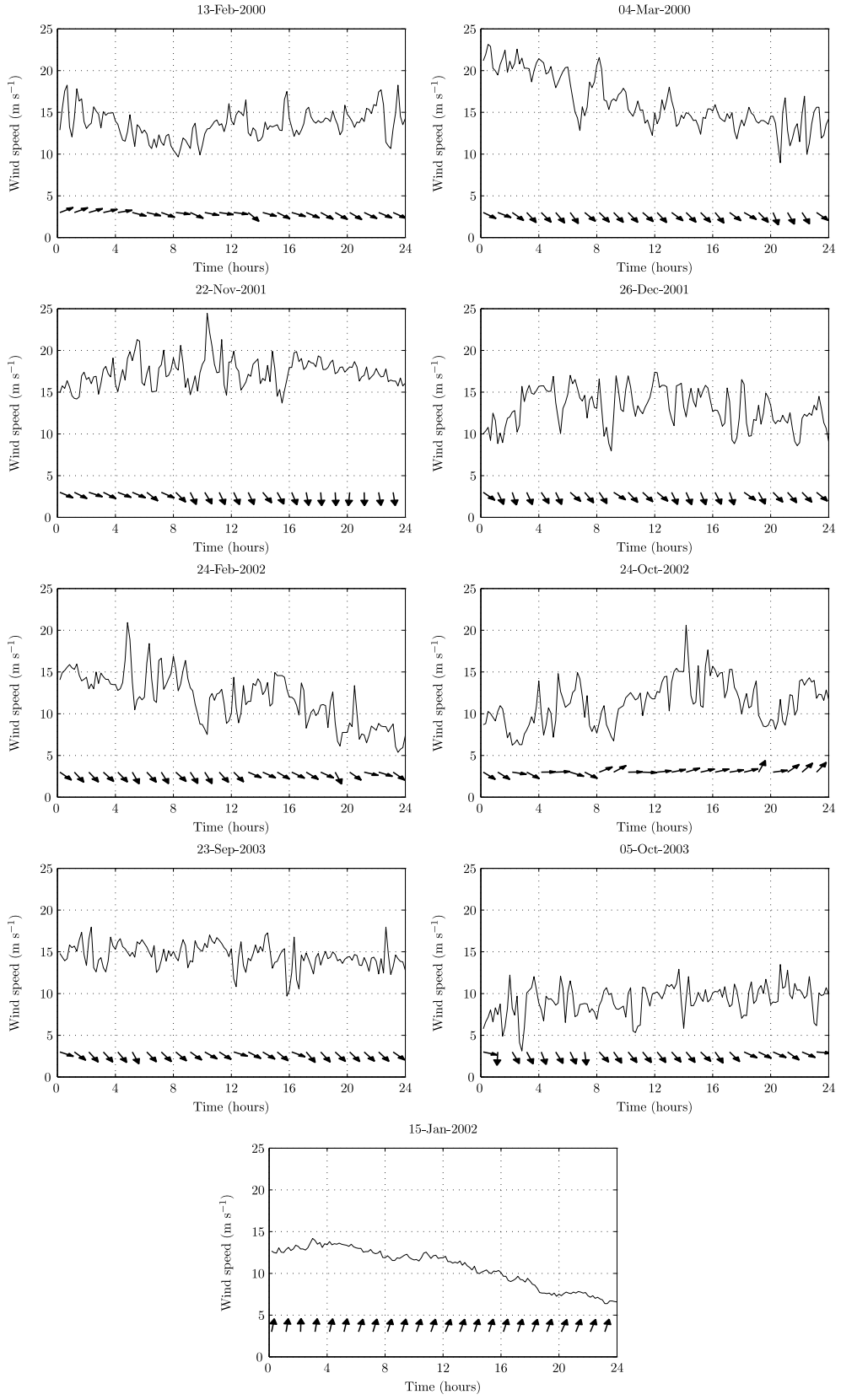


Figure 8.1: Observed wind speed at a height of 62 m (lines) and wind direction at a height of 43 m (arrows) for the cases chosen for the WRF simulations. The final case is the control case of very low amplitude mesoscale variability.

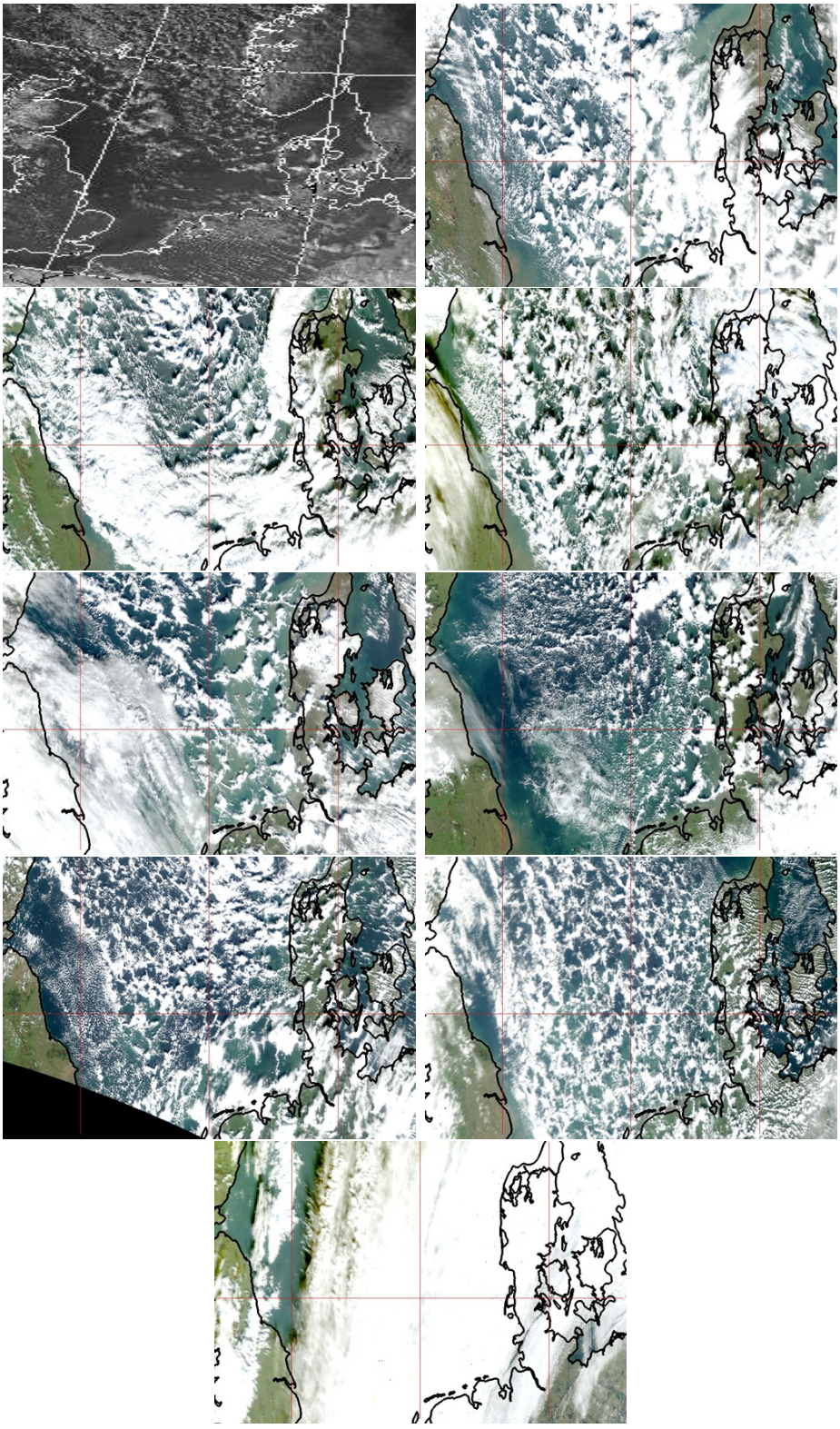


Figure 8.2: Satellite pictures for the nine cases shown in figure 8.1. In each case, the satellite picture is valid at about 12 UTC, which is 12 hours after the beginning of the time series in figure 8.1. The first picture is from the NOAA14 Satellite, and was sourced from http://www.sat.dundee.ac.uk, while the remaining pictures are from the TERRA and AQUA MODIS satellites, sourced from http://ladsweb.nascom.nasa.gov/

8.4 Methodology

8.4.1 Model setup

The nine cases were simulated using 4 nested domains with horizontal grid spacings of 54 km, 18 km, 6 km and 2 km respectively. The four domains are illustrated in figure 8.3. The setup of the WRF model (ARW Version 3.2) is described in table 8.1. There were 37 vertical levels, with the lowest five mass levels close to 14 m, 53 m, 105 m, 163 m and 228 m. These heights change slightly with time, since WRF uses a terrain-following hydrostatic-presssure vertical coordinate. However, domain 4 mainly consists of the relatively flat Danish landscape and the North Sea, so the heights will not change much with time.

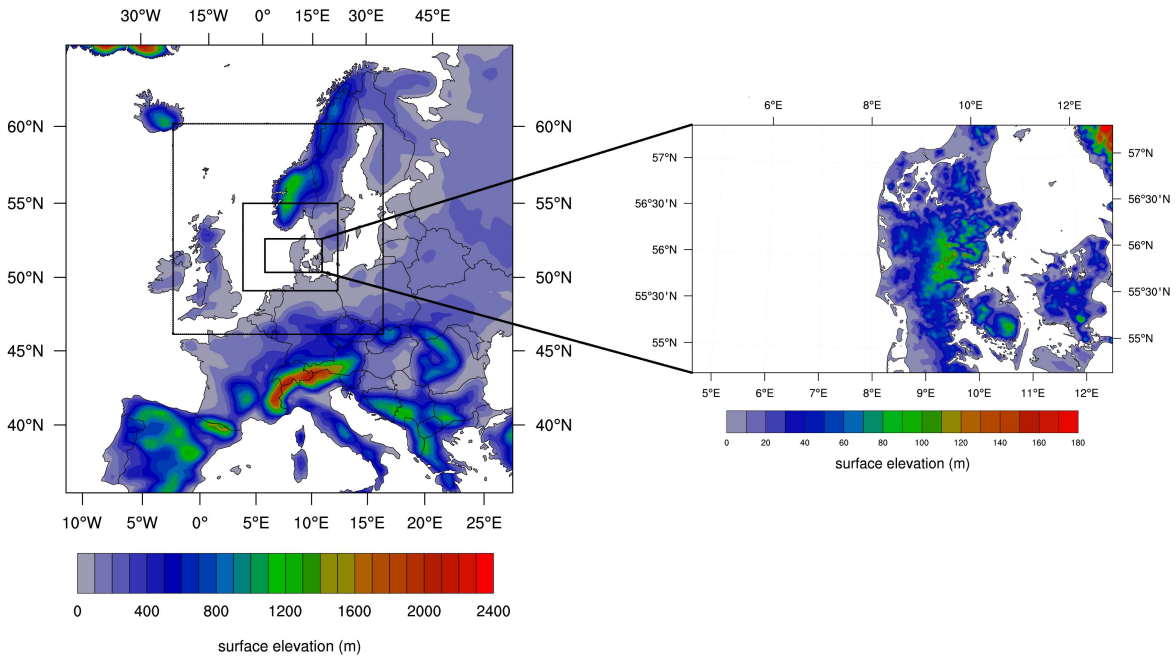


Figure 8.3: The four nested domains for the WRF simulation. Domain 4 is shown in greater detail, with a finer scale for the surface elevation contours. Note the different colour scale for the two plots.

8.4.2 Intra-hourly wind speed variance

Although the smallest domain of the model was run with a time step of 12 s, the model output was only saved once per hour, due to the prohibitive disk storage requirements for more frequent model output. The wind speed was stored for every time step at a few locations for bottom five model levels using functionality that is supplied with WRF. This does not, however, give a spatial representation of the intensity of the temporal wind fluctuations. This was partly addressed by implementing a single-pass algorithm in WRF to calculate the variance of the wind speed at the end of each hour. It was important that it was a single-pass algorithm, because this avoided storing all values until the end of the hour. The single-pass variance algorithm (Welford, 1961) is based on updating the mean, $\bar{x}_n$, and the sum of squares, P_n, as each new value of the time series, x_n, becomes available. The variance

Domain	1	2	3	4
dx	54 km	27 km	6 km	2 km
dt	216 s	108 s	36 s	12 s
Topography resolution	$10'$	$10'$	$2'$	$30''$
Cumulus parametrisation	K-F	K-F	K-F	None
Vertical levels	37			
Microphysics scheme	Thompson et al scheme (Thompson et al., 2004)			
PBL physics	MYNN scheme (Nakanishi and Niino, 2009)			
Long-wave radiation physics	RRTM scheme (Mlawer et al., 1997)			
Short-wave radiation physics	Dudhia scheme (Dudhia, 1989)			
6th order diffusion	On			
Diffusion	Diffusion along coordinate surfaces			

Table 8.1: Setup of the WRF model that was used in the nine case study simulations. K-F stands for the Kain-Fritsch cumulus scheme (Kain and Fritsch, 1990; Skamarock et al., 2008).

is then calculated by dividing the sum of squares by the number of values in the time series. The algorithm is described in equations 8.1, 8.2 and 8.3. A derivation of the equation for the updating sum of squares is given in appendix A.

Updating mean

$$\bar{x}_n = \bar{x}_{n-1} + \frac{1}{n}(x_n - \bar{x}_{n-1}) \tag{8.1}$$

Updating sum of squares

$$\begin{aligned} P_n &= \sum_{i=1}^{n}(x_i - \bar{x}_n)^2 \\ &= P_{n-1} + (x_n - \bar{x}_{n-1})(x_n - \bar{x}_n) \end{aligned} \tag{8.2}$$

Updating Variance

$$\sigma_n^2 = \frac{P_n}{n-1} \tag{8.3}$$

The updating variance was implemented in WRF in as a function that was called at every time step. At the end of each hour, the variance was written to the WRF history files, and the updating mean and sum of squares were reset to zero. The variance does not give any information about the frequencies of the wind variability, and is therefore only of limited use. Other single pass statistics that could be implemented in a similar way would be a maximum ten-minute wind speed change, or the integrated wind energy potential over each hour.

8.5 Results

Time series of model values for the point closest to the M2 meteorological mast at Horns Rev I were saved for the third model level for each of the nine cases presented in figure 8.1.

The model output, which was stored at every 12 s time step, was averaged to match the 10 minute time step of the observations. Each simulation started 12 hours before the 24 hour periods of interest, and ran for 36 hours to cover the whole case. This means the first few hours of the simulation, which are initialised from the coarse resolution initial conditions, can be disregarded. The full 36 hour forecasts at a height of 53 m and observed time series at a height of 62 m are shown in figure 8.4. In all cases apart from case 4 (26 December 2001), and the control case, large wind fluctuations are apparent in both the observed and modelled time series.

The MSLP for domain 2 is shown in figure 8.5 for each of the 8 severe variability cases and the case of low wind variability. For the 8 cases of severe wind variability, the MSLP patterns correspond closely to the synoptic types 15, 16 and 22 that were identified in chapter 5 as high-risk situations for severe wind variability. All have a deep northerly component to the flow.

To diagnose the variability in the forecast and observed time series in a simple way, the variance was calculated over a three hour moving window for both the observed and the modelled time series. The variance does not reflect the frequencies at which the fluctuations occur, and is therefore responds strongly to a single large change in wind speed (such as at the end of the case on 25 December 2001) as well as to extended periods of smaller fluctuations. However, it is a useful diagnostic. Another tool that could have been used to assess the modelled wind variability is the Hilbert-Huang transform that was used for the analysis of variability in observed time series in chapters 3 and 4. However, for such short time series, the end effects present an unresolved problem. There are interesting possibilities of using the Hilbert-Huang transform for scale selective model verification, as discussed in Vincent et al. (2009), although the right way to deal with end effects is far from clear at this stage.

For example, in the case on 21 November 2001, the serious model error in the mean wind at hour 6 does not actually appear in the time series of variance. In this case, it is possible that the mean wind speed in the model was incorrect from the outset due to wrong initial conditions, but that the right kind of physical processes to cause large wind fluctuations developed anyway. For the case on 22 September 2003, the forecast looks rather good in the first 12 hours in terms of the mean wind speed, but the time series of variance reveal a serious deficiency in variability. Together, the time series demonstrate the different kind of errors that can occur in models of wind speed. Even the control case on 14 January 2002, for which model correctly simulates the very low mesoscale variability, there is a phase error of several hours in the timing of the maximum wind speed.

A serious model error in both the mean wind speed and variability occurred on 25 December 2001. Close inspection of the wind and pressure fields for the simulation of this case reveal the development of a local high pressure ridge over western Denmark and a low pressure trough over the North Sea just to the west of Denmark. The configuration of the ridge and trough resulted in light wind speeds over Horns Rev, and an easterly instead of northwesterly wind direction, as seen in the small wave in the MSLP contour over western Denmark in figure 8.5. Further westwards in the North Sea, strong north-northwesterly flow and highly variable cellular patterns persisted in the model. Interestingly, the local trough and ridge dissipated towards the end of the 36 hour simulation, allowing Horns Rev to be influenced again by the north-northwesterly flow. This is reflected in the improvement in the simulated wind speeds for the last couple of hours of the simulation. The errors in this case are extremely

interesting in themselves, and might relate to errors in the initial conditions. This highlights the fact that if the large scale flow is not correct, then there is no possibility of developing correct mesoscale features.

Vertical velocity for the third model level is shown in figure 8.6 for each case for hour 18 of simulation. In every case, apart from the control simulation, cellular structures with a similar length scale to those in the satellite pictures developed. The cellular patterns did not generally cover the whole North Sea area in the way that they did in most of the satellite pictures. Instead they tended to form and decay in different parts of the domain and different times. This could partly be due to limitations imposed by the domain size.

The updating hourly variance, as described in equations 8.1 to 8.3 for the third model level is shown for each case for hour 18 of the simulation. It is clear that the regions of enhanced updating hourly variance correspond closely to the regions of cellular convection, despite the fact that the updating hourly variance is necessarily responding to sub-hour scale variance, while the time scale of the fluctuations associated with the cellular convection is greater than one hour in most cases. The updating hourly variance therefore presents a useful way of quantifying the existence of mesoscale fluctuations in the model.

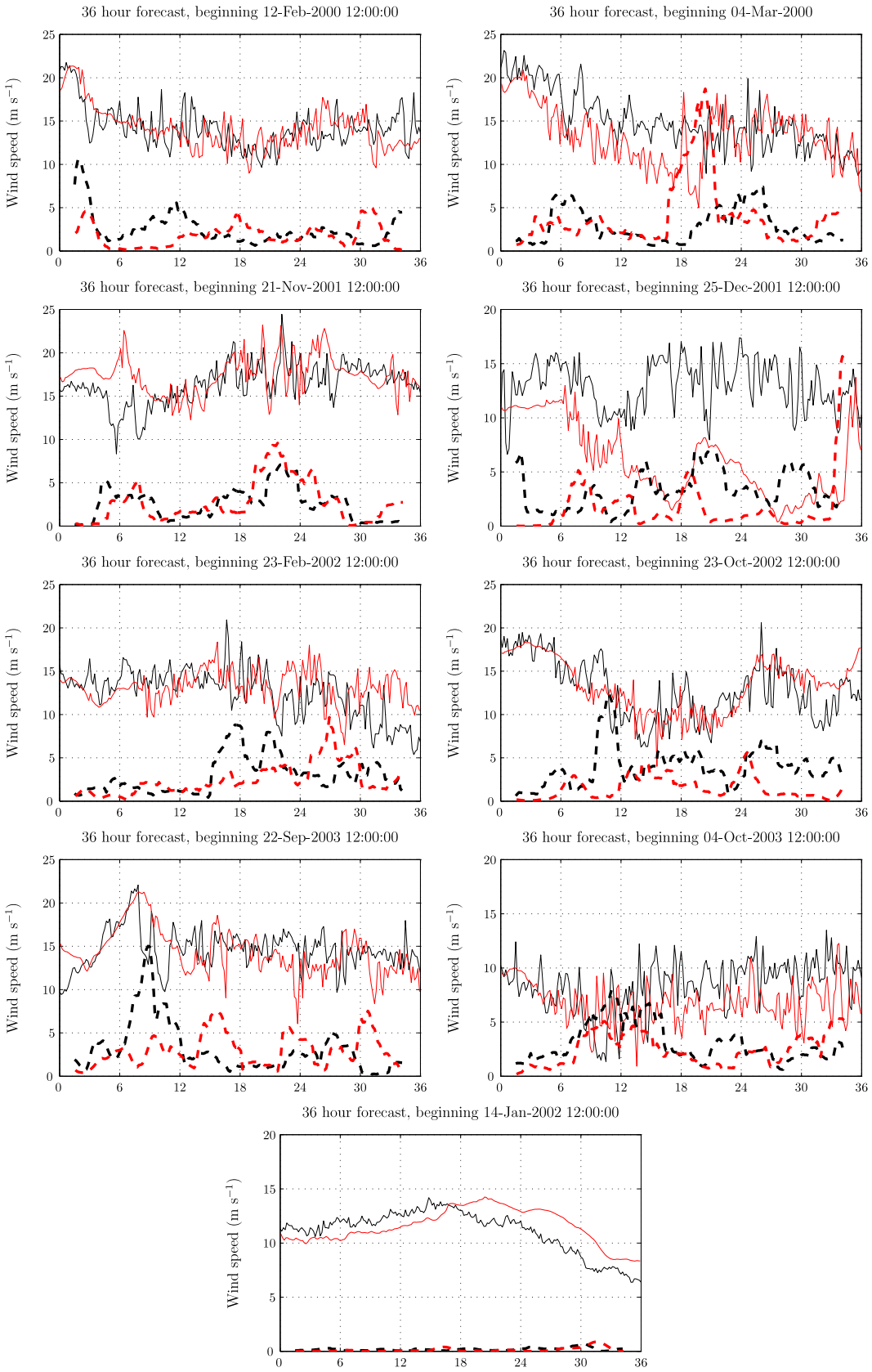


Figure 8.4: Simulated and observed time series at Horns Rev I for the nine cases shown in figure 8.1. Black - Observations at a height of 62 m. Red - Model simulation at a height of 53 m. Solid lines - wind speed. Dashed lines - three hour moving variance. The final case is the control case of very low amplitude mesoscale variability.

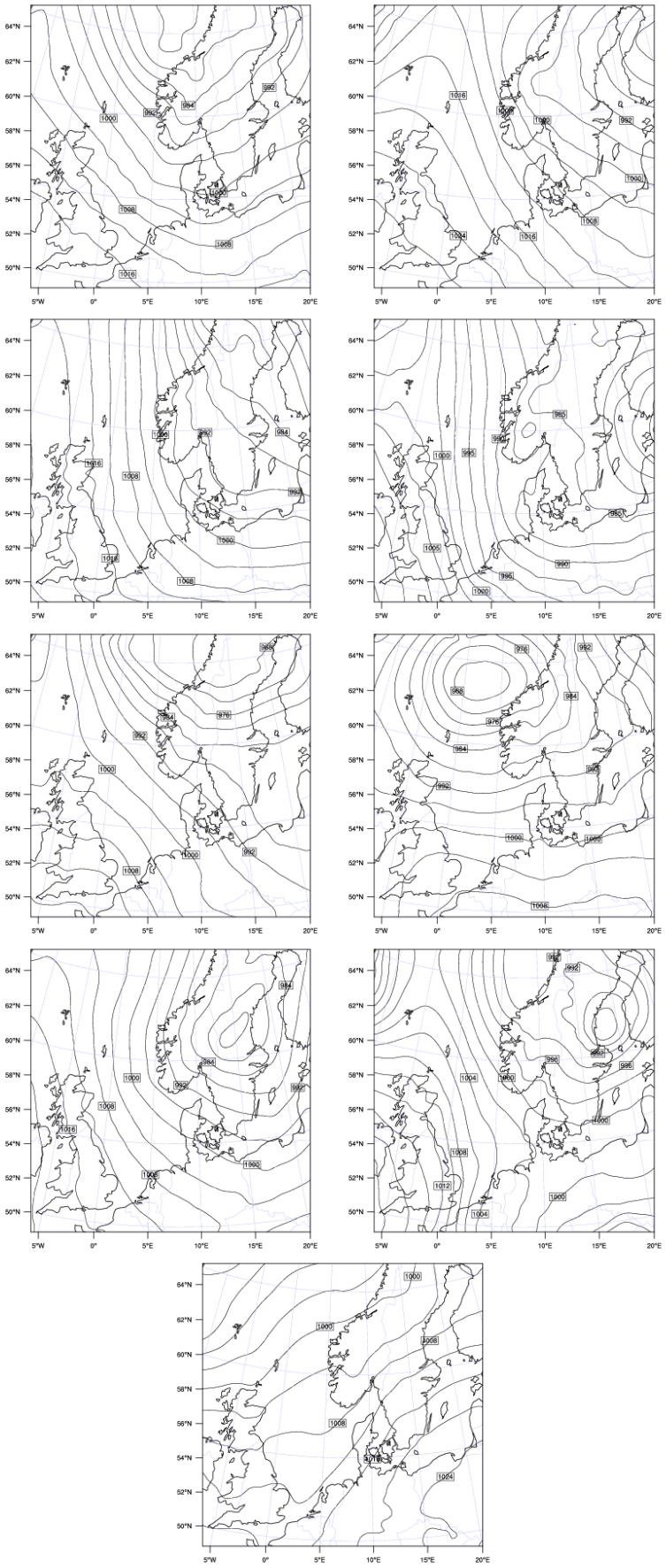


Figure 8.5: Mean sea level pressure for domain 2 for the nine cases. The plots correspond to the 18th hour of the 36 hour simulation period.

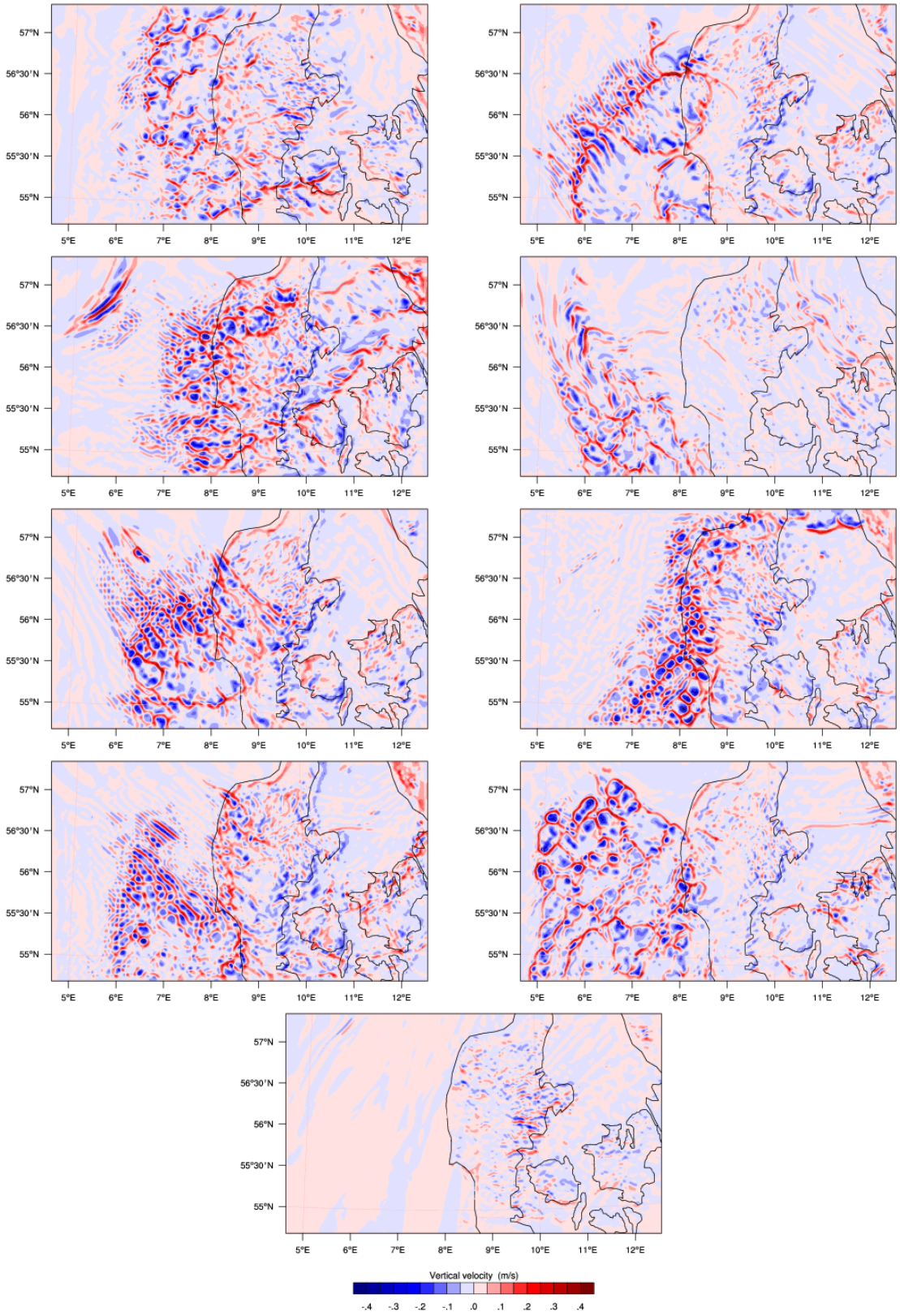


Figure 8.6: Vertical velocity for the 3rd model level (105 m) for the 9WRF simulations. The final case is the control case of very low amplitude mesoscale variability. The plots are for the 18th hour of the 36 hour simulation.

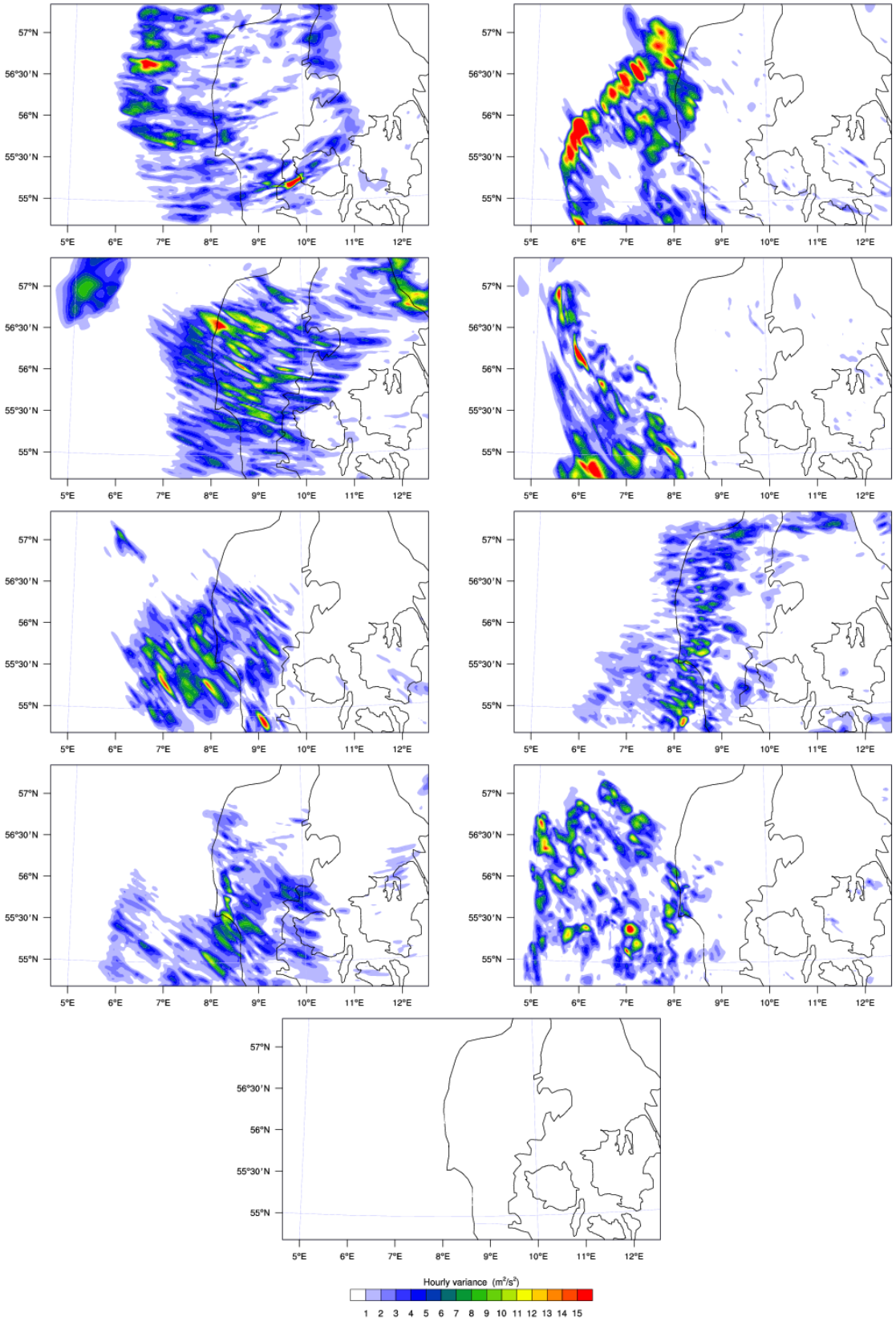


Figure 8.7: Updating hourly variance for the 3rd model level (105 m) for the 8 cases, and the ninth control case of very low amplitude mesoscale variability. The plots are for the 18th hour of the 36 hour simulation.

8.6 Discussion and Conclusions

In this chapter, a set of 8 days where large amplitude wind variability was observed at the Horns Rev I wind farm were simulated using the WRF model, and the results were analysed according to modelled time series, vertical velocity, and a special algorithm for updating hourly variance that was implemented in the WRF model. All 8 cases fitted into the synoptic patterns that were identified as having a high risk for large amplitude wind variability in chapter 5, and OCC was evident in the satellite picture for all the cases. A ninth case where very little hour-scale wind variability was observed was also modelled as a control case.

It was found that in all 8 cases, cellular-like patterns emerged in the vertical velocity. The cellular patterns differed in their diameter and exact configuration, but common to all cases was the fact that the cells were predominantly over the sea, and weakened or disappeared altogether over western Denmark. The exact coverage of OCC was not well modelled, although the results in figure 8.6 suggest that the OCC might be restricted from developing in the upstream corner of the domain (since we do not expect fully formed cells to advect in from outside the domain), and this limits the area over which the cells could form. However, during the 36 hour simulations, areas of OCC were constantly forming and decaying over different parts of the domain, and the results presented here are just an instantaneous picture of their development.

As well as the cellular patterns which are apparent in the spatial fields, large wind fluctuations are seen in the modelled time-series that closely resemble the observed time series. The implications of this are two-fold. Firstly, from a practical point of view it means that time series forecasts from high resolution mesoscale models can provide useful and realistic forecasts of the mesoscale variability for applications such as wind power forecasting, provided the large scale weather patterns are correct. Secondly, from a modelling point of view, it means that the energy that is transformed into mesoscale wind fluctuations in the presence of OCC is realistic.

The updating hourly variance was an interesting extension to the output that is normally available from a mesoscale model. NWP output is often provided to end-users at hourly or even longer intervals, usually with the justification that the expense of saving and transferring higher frequency data is prohibitive. For example, wind forecasting software such as the WPPT (Nielsen, 1999) often receive hourly input from NWP data. The updating hourly variance means that not all higher resolution information is lost from the model run, and provides some retrospective insight into what goes on in-between hourly model history files. Unfortunately, the variance does not give any information about the frequency of the fluctuations that were occurring during the hour, and different spectral behaviour could give the same variance. Nonetheless, the updating hourly variance could provide a useful addition to NWP output.

Actually, any statistic that can be calculated using a single pass methodology could be easily implemented in the same frame-work that was used for calculating the updating hourly variance. Examples of statistics that could be implemented are maximum and minimum hourly wind speed, maximum 10 minute wind change or the integrated wind energy production. Of course, an ideal statistic would be an updating set of Fourier coefficients for the mesoscale range of the wind speed spectrum, but the implementation of this as a single pass operation is far from obvious.

The case-study framework was useful for the detailed analysis of the model's response to a few selected weather conditions. However, this is far from ideal if one wishes to statistically verify the model in a long term sense, and it is definitely inadequate if one wishes to compare different models or different versions of a model. Ideally, a long time series of model runs would have been analysed (preferably at least one of the four years 2000–2003 for which wind speed data was available at Horns Rev). This was not possible due to various time and computing constraints, but is important future work. Long term and statistically rigorous verification of mesoscale variability raises the interesting topic of the development of verification statistics for fluctuations that are almost certain to be out-of-phase, even if they are physically and statistically correct.

Chapter 9

A simplified WRF model of open cellular convection

9.1 Introduction

The results presented in chapter 8 suggested that the WRF model with horizontal grid spacing of 2 km is able to simulate some mesoscale features that have length scales of tens of kilometers. In particular, several cases of open cellular convection over the North Sea were simulated in such a way that hour-scale variability and realistic cellular convection patterns were generated.

The case-study approach of chapter 8 is limited by the difficulty of separating the many factors that control mesoscale forcing in a numerical weather prediction model. For example, variability associated with isolated events such as fronts and the effects of topography may be spectrally blended with variability from patterns of mesoscale convection, and the constantly changing properties of the air ensure that fields are non-stationary in both space and time. These limitations are partially addressed in this chapter by running the WRF model with a simplified setup, with a specific focus on simulating open cellular convection (OCC). The simplified model has no topography, no large scale weather forcing and no surface inhomogeneities.

In the first part of this chapter, the technical setup of the simplified WRF model is described, and the results of changing various model settings are presented. Specific changes that were made to the WRF code to implement the simplified model will also be explained. In the second part of the chapter, the sensitivity of the modelled OCC to small changes in the environmental forcing is investigated.

9.2 The idealised WRF simulation

The simplified WRF model is based on the idealised sea breeze circulation and large eddy simulation (similar to that described in Moeng et al. (2007)) that are provided with the basic WRF package. The model domain is three dimensional, with periodic boundary conditions in the x and y directions. The lower boundary is a water surface with a fixed, specified sea surface temperature, and the model is initialised by vertical profiles of the two wind components, potential temperature and water vapour mixing ratio. A random perturbation of amplitude 0.05 K was added to the temperature field to initialise the convection.

The simulations are based on WRF ARW version 3.0, with several minor changes were to the existing initialisation scripts for idealised experiments. They are:

1. The sea surface temperature was added as a registry variable that could be controlled at run-time rather than at compile time.
2. The specification of vertical levels was changed so that it was the same as in the full WRF model, rather than the existing simplified way of specifying levels.
3. The time series output routine was changed so that it took an ij grid point instead of a latitude–longitude point.
4. The Coriolis parameter was set as the relevant value for 55°N.

A control experiment was defined as a reference for testing various aspects of the model setup. The specifications of the control experiment are described in table 9.1.

Parameter	Value
dx	1 km
Number of grid points, x dir	349
Number of grid points, y dir	349
dt	6 s
Cumulus parametrisation	None
Vertical levels	37
Microphysics scheme	Thompson et al scheme (Thompson et al., 2004)
PBL physics	YSU scheme (Hong et al., 2006)
Long-wave radiation physics	RRTM scheme (Mlawer et al., 1997)
6th order diffusion	On
Diffusion	Diffusion along coordinate surfaces
Vertical advection scheme	3rd order Runge Kutta
Horizontal advection scheme	5th order Runge Kutta

Table 9.1: Set up of the WRF model for the control case of idealised open cellular convection simulations

Vertical profiles of the wind components, potential temperature and water vapour mixing ratio were extracted from the real WRF simulation on 24 October 2002 (case 6 in chapter 8) and used to initialise the idealised experiment. An alternative to using a simulated profile

from WRF would have been to use real observations from a radiosonde. The simulated WRF profile was used because it is representative of balanced model conditions, and because the goal here was to simplify the mesoscale structures that had already been simulated in the model. The profile that was used to initialise the model is shown in figure 9.1. For comparison, an observed profile from the Ekofisk oil platform for the same time is included on the plot. Although the Ekofisk oil platform is around 300 km further west than Horns Rev, the satellite picture in figure 9.2 suggests that the meteorological conditions were similar at the Horns Rev and Ekofisk sites, and that the WRF profile is therefore broadly representative of the actual atmospheric conditions. The main difference between the simulated and observed profiles is that the observed profile is drier than the simulated profile.

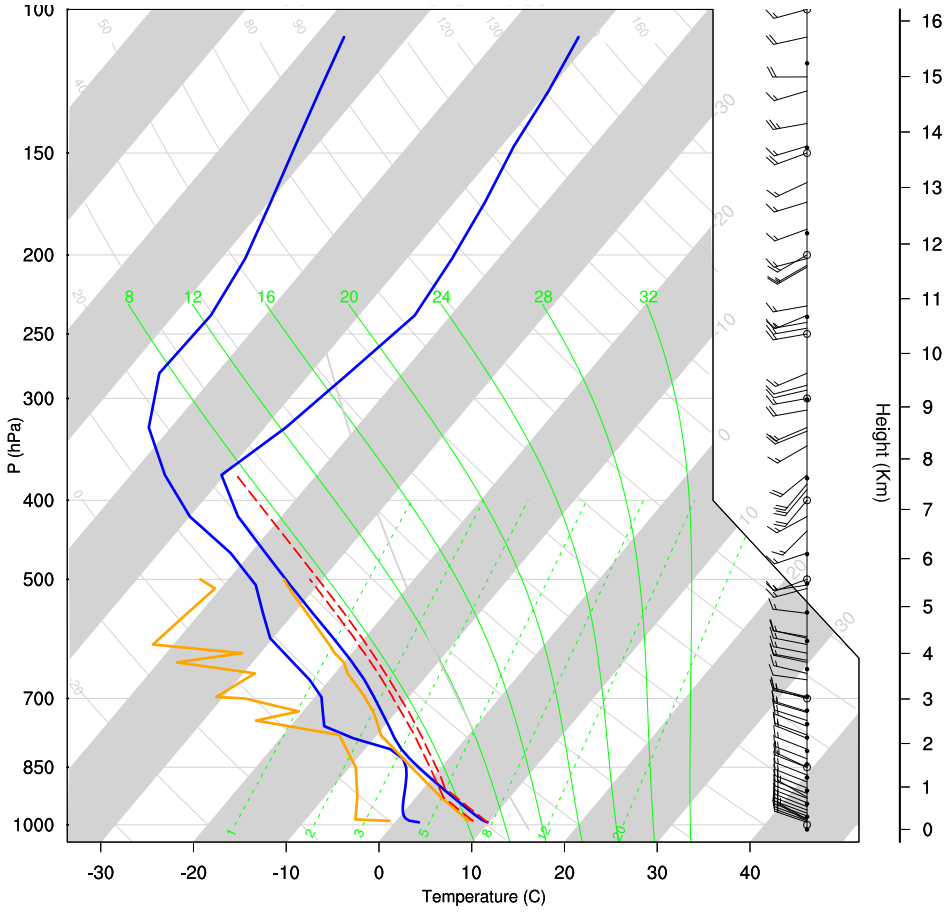


Figure 9.1: Simulated profile of temperature, dewpoint temperature (blue lines) and wind barbs (right panel) from the real WRF simulation initialised on 23 October 2002 12 UTC, for simulation hour 12 (24 October 2002 00 UTC). The observed temperature and dewpoint temperature profiles from a radiosonde released from the Ekofisk oil platform are shown in Orange. Dashed red lines show the path of parcels of surface air lifted adiabatically. Solid green lines: saturated adiabiats. Solid light grey lines: dry adiabats. Wide grey stripes: constant potential temperature lines. Dotted green lines: constant mixing ratio lines.

The control simulation ran for 12 hours, and started at 1800 UTC so that there was no incoming solar radiation during the simulation. The vertical velocity for hours 2, 5 and 12 of the simulation is shown in figure 9.3. By the end of the second simulation hour, a regular pattern of convection cells had developed. The cells had a diameter of 2–5 grid points (2–5 km) and were formed in a pattern of regular lines, suggesting that the convection at this

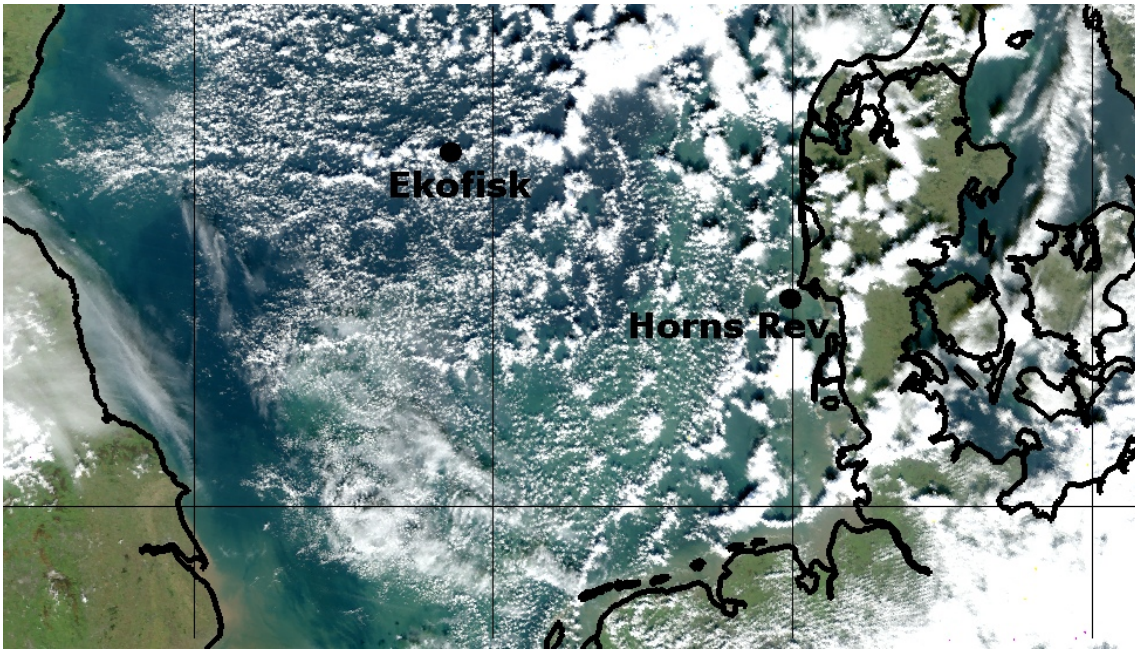


Figure 9.2: Satellite picture of the North Sea on 24 October 2002 11 UTC, showing the location of the Horns Rev wind farm and the Ekofisk oil platform.

stage was modulated by the model grid itself.

The cells grew in diameter and intensity during the subsequent hours of the simulation, until at hours 5 and 6 they were well formed open cells, with upwards velocity at the cell walls and downwards vertical velocity in the cell centres. The cells had a vertical extent of around that of the boundary layer height, although strong up-drafts extended above the boundary layer at some of the cell walls. After the 6th simulation hour, the cells continued to grow in size, and simultaneously weaken in intensity, until by the end of the 12 hour simulation period, the domain was covered in a network of thin lines of upwards vertical velocity separated by wide areas of mostly downwards vertical velocity, but without any obvious arrangement into a regular cellular structure.

One of the ways in which the idealised simulation is different to a mesoscale model running as a simulation of the real weather is that there is no regeneration of the atmospheric conditions, since the horizontal boundary conditions are periodic, and the unstable conditions and air-sea temperature difference required for the maintenance of the OCC are eroded by the OCC itself. There is no particular instant when the OCC can be considered properly developed, since it will already have begun to modify its atmosphere from the very first time steps. However, during hours 5 and 6 of the simulation, the cells appeared to have developed into well defined cells. Before this, they qualitatively appeared to be still growing, and after this they began to decay. Therefore, these hours were considered as the interesting hours of the simulation for subsequent analysis.

The surface divergence caused by the large variation in vertical velocity near the cell walls was necessarily associated with variation in the horizontal wind speed at the surface as well. In figure 9.4, the horizontal wind speed is shown in shaded contours, with vertical wind speed shown as superimposed contour lines. Here it is seen that fluctuations in horizontal wind

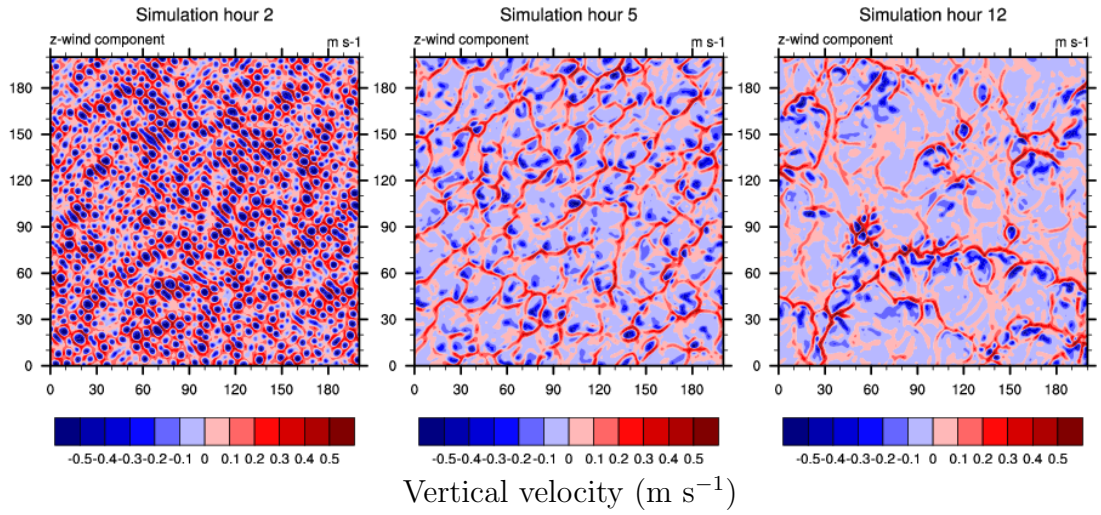


Figure 9.3: Vertical velocity (third model level) for hours 2, 5 and 12 of the control simulation. The axes are labelled in number of grid points, and the horizontal grid spacing of the model is 1 km.

speed of up to $\pm$ 3 m s^{-1} occur, which is consistent with the magnitude of fluctuations that were observed and simulated at Horns Rev. The figure is not only a fascinating picture of the complex patterns developed by the mesoscale model, but a clear demonstration of why OCC and other mesoscale effects are of critical relevance to large wind power installations over the sea. Figure 9.4 gives some hint of a systematic asymmetry in the cells, with an area of stronger convergence and stronger vertical velocity on the southwest side of the cells. This tendency is also mentioned in the context of an observational study by Sikora et al. (2010), who identified an arc of large cumulus clouds on around the downwind edge of cells in satellite pictures, and connected this with a squall of stronger wind speeds. It is not known whether the asymmetry of the modelled cells is physical, or an unrealistic feature of the model simulation, and this question calls for further investigation.

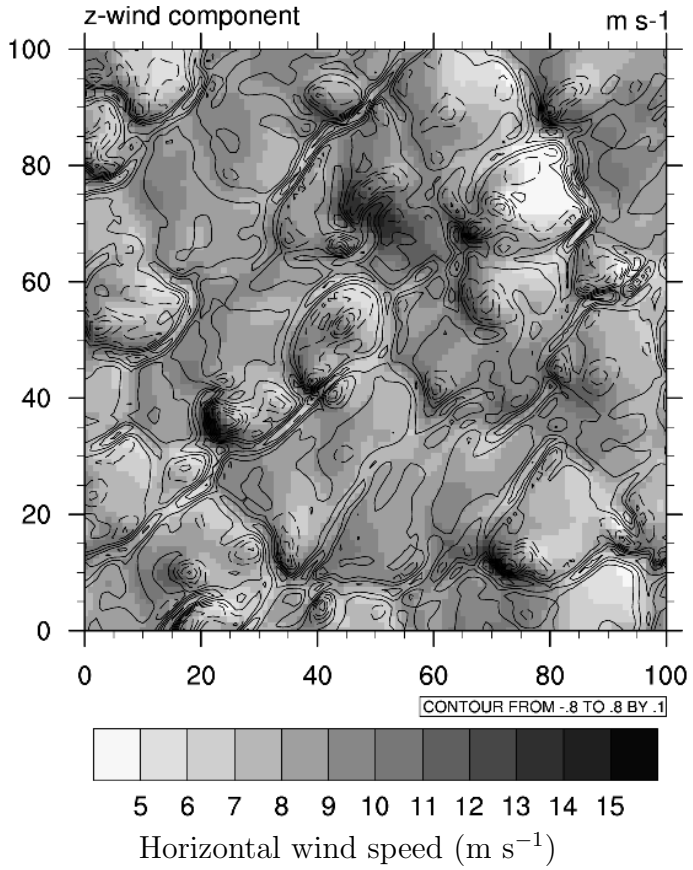


Figure 9.4: Horizontal wind speed (shaded contours) and vertical velocity (contour lines) for the 6th simulation hour.

9.3 Definition of model parameters

There are many combinations of parameters which must be chosen to specify the dynamical and physical schemes that are used in WRF. This is also a problem for real WRF simulations, but running the model in an idealised setting provides a unique opportunity to isolate the effect of specific changes to the model.

Dynamical and physical aspects of the model setup were tested in a series of 9 experiments, which are described in table 9.2. Each set of parameters was tested and compared to the control simulation, which is experiment number 1 in table 9.2. In each case, only one parameter was modified at a time which was meant to isolate the effect of that parameter. This neglects the fact that the model settings do not operate in isolation, and may have a different effects when used in particular combinations. Results were analysed by inspection of vertical velocity fields, surface divergence and calculation of the spatial spectrum of vertical velocity.

Two dimensional spectra of the idealised model fields were calculated to assess the impact on the cell size and spectral power at high wavenumbers. Following Errico (1985), the two dimensional spectrum, which is a function of wavenumber in the i and j directions, was mapped to one dimension by summing the variance within annuli of constant wavenumber

radius. The periodic boundary conditions mean that no detrending of the fields was required. Since the fields of the idealised experiment were statistically stationary in the spatial dimension, the two dimensional spectra at a fixed time gave a more reliable result than a spectrum of the time series, which would be non-stationary due to the evolving properties of the atmosphere and the gradual growth of the cells. The spatial spectra also have the advantage of good data representativity, which would not have been the case if a simple one dimensional spectrum was calculated based on a cross section through the domain.

Experiment	1	2	3	4	5	6	7	8	9
dx (km)	1	0.7	2	1	1	1	1	1	1
6th order diffusion	1	1	1	0	2	1	1	1	1
PD moisture advection	T	T	T	T	T	F	T	T	T
PD tke advection	F	F	F	F	F	F	F	F	F
Coriolis	1	1	1	1	1	1	P	0	1
PBL scheme	Y	Y	Y	Y	Y	Y	Y	Y	M

Table 9.2: Experiments to test the model parameters. T and F refer to TRUE and FALSE switches, while the meanings of the other values are discussed in the text. The first column (shaded) is the control experiment.

The results of the experiments in table 9.2 are used define a revised model setup that is treated as a control case for a series of experiments that test the sensitivity of simulated cells to small changes in the atmospheric conditions in section 9.4.

9.3.1 Horizontal grid spacing

The sensitivity of the simulation to horizontal grid spacing was tested for the values of 2km, 1 km and 700 m. In all cases, vertical diffusion was taken into account by the boundary layer parameterisation. The three horizontal grid spacings were chosen so that they were large enough that the vertical diffusion could be taken into account by the boundary layer parameterisation, but small enough that a cumulus parameterisation was not required. As argued by Molinari and Dudek (1992), a mesoscale model with a horizontal grid spacing less than about 5 km should be able to explicitly resolve cumulus convection, so by limiting the grid spacing of the experiments to 2 km, any ambiguity as to whether a cumulus parameterisation should be employed is avoided. As demonstrated by Skamarock (2004), the spectral power of features smaller than about seven times the horizontal grid spacing starts to be attenuated in comparison to observed spectra. This means that cells with a diameter of around 20 km should be close to the limit of predictability for a horizontal grid spacing of 2 km. The vertical velocity at the end of the 6th simulation hour is shown for the three cases in figure 9.5, where a square sample of the domain of size 200 km^2 is shown in each case.

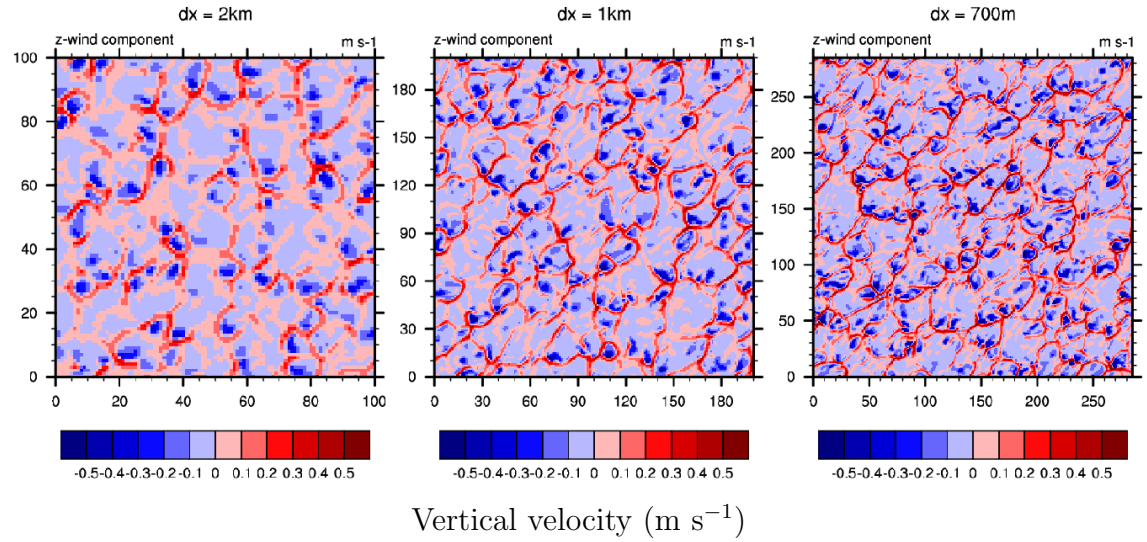


Figure 9.5: Experiments 2, 1 (control) and 3 respectively. Vertical velocity (third model level) for 2 km (left), 1 km (centre) and 700 m horizontal grid spacing. The axes are labelled in number of grid points, so that all three plots show an area of 200 km^2. No smoothing is applied, so the pixelated appearance shows individual grid squares.

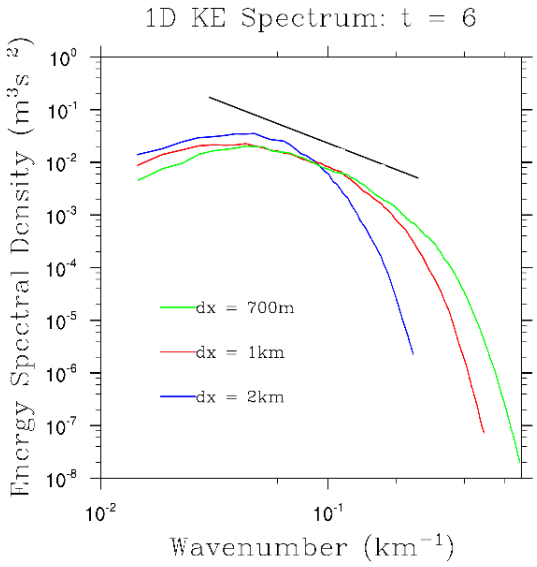


Figure 9.6: Kinetic energy spectra for experiments 2, 1 (control) and 3 respectively, for the third model level at the end of the 6th simulation hour. The black line shows a spectral slope of $-5/3$.

The results show that the overall structure and scale of the OCC is similar for the three experiments, but that the cells are generally poorly defined and less intense in the 2 km experiment. The kinetic energy spectrum (figure 9.6) indicates that the peak of the spectrum (corresponding to cell size) is similar for the 1 km and 700 m experiments, and that the spectra have a similar shape up to a wavenumber of about 0.12, corresponding to a wavelength of about 8 km. Both the 1 km and the 700 m experiments seem to briefly reach a spectral slope of $-5/3$ before they begin to decay with a steeper gradient. The similarity

between the 1 km and 700 m experiments suggested that the 1 km grid spacing was actually sufficient for modelling the OCC in this case.

9.3.2 6th order diffusion

6th order diffusion was introduced into the WRF model to address the issue of insufficient implicit diffusion, particularly in low-wind situations where the implicit diffusion arising out of the odd-ordered advection schemes is small (Knievel et al., 2007). Since the scheme is 6th order, it is scale selective and designed to act on features that are smaller than about six times the horizontal grid spacing of the model. An extension to the scheme also ensures monotonicity and limits counter-gradient fluxes. Knievel et al. (2007) demonstrated the differences between simulations with and without 6th order diffusion, and with and without the flux limiter, and showed that in the simulation with 6th order diffusion, the patterns in surface divergence were less aligned with the model grid and more induced by the terrain. The flux limiter resulted in slightly smaller extrema compared to the case without the flux limiter, although it introduced the disadvantages of less scale selectivity and slightly less diffusion at the highest wavenumbers. Here, three settings for the 6th order diffusion are tested: No 6th order diffusion (0), 6th order diffusion without the flux limiter (1), and 6th order diffusion with the flux limiter (2) (experiments 4, 5 and 1 in table 9.2).

Plots of vertical velocity and surface divergence for the three experiments are shown in figure 9.7. The plots indicate stronger upwards vertical velocity and a more complicated cell structure for the case without 6th order diffusion. Since the greatest surface divergence occurs in the vicinity of the cell walls, the more complex cell structure results in greater noise in the surface divergence. In particular, areas of positive and negative surface divergence are found immediately adjacent to one another, while the surface divergence for the cases with 6th order diffusion are cleaner. The 6th order diffusion with flux limiter is dominated almost exclusively by the cells themselves. Figure 9.8 shows the kinetic energy spectra for the three experiments. The spectra show that, as expected, the two experiments that had 6th order diffusion have less spectral energy at wavelengths greater than about 6 times the horizontal grid spacing. However, there are also differences in the spectra for wavenumbers longer than $6\Delta x$, which means that the filter is not perfectly scale selective in this case. The difference is especially noticeable for the case with the flux limiter, where the peak of the spectrum actually seems to be shifted to slightly lower wavenumbers. As discussed by Knievel et al. (2007), these differences could come about because in moist convection, the growth of individual cells is dependent on certain thresholds being exceeded, so that damping extrema on small scales could feedback onto larger scales. Despite the uncertainty about whether to use the flux limiter, it was decided to include it in the revised model setup because the vertical velocity fields and surface divergence were clearly dominated almost exclusively by the cells themselves, with very little higher frequency noise.

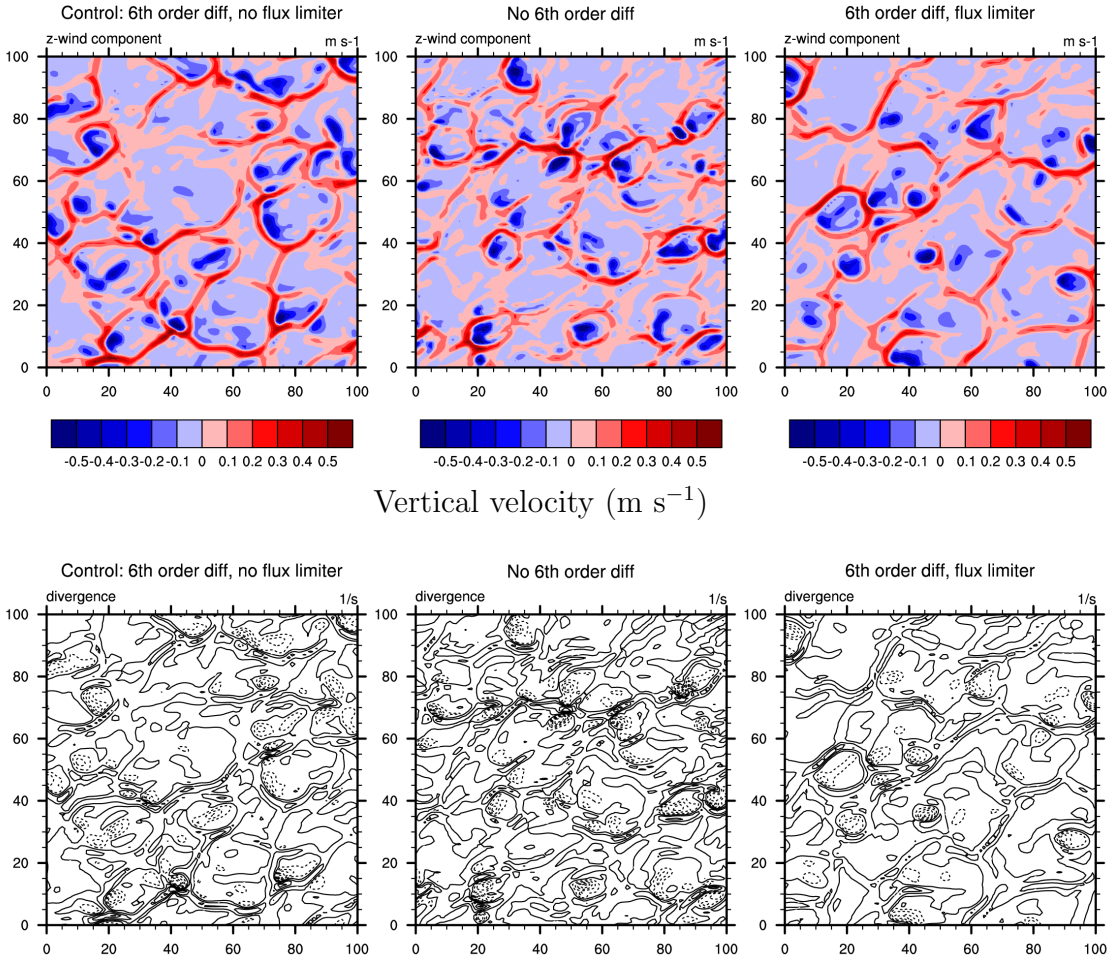


Figure 9.7: Experiments 1 (control), 4 and 5 respectively. Vertical velocity for the third model level (upper panels) for 6th order diffusion without flux limiter (left), no 6th order diffusion (centre) and 6th order diffusion with flux limiter (right). Divergence at the lowest model level (lower panels) for the corresponding experiments (positive: solid, negative: dashed). Divergence is contoured every 2×10^5 s^{-1}.

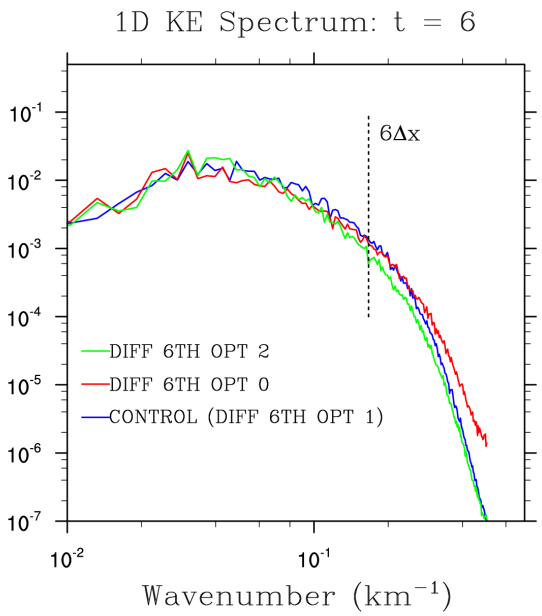


Figure 9.8: Kinetic energy spectra for experiments 1 (control), 4 and 5 respectively, for the third model level at the end of the 6th simulation hour. The vertical dashed line indicates the wavenumber corresponding to a wavelength of six times the horizontal grid spacing of the model.

9.3.3 Planetary Boundary Layer scheme

The planetary boundary layer (PBL) scheme is responsible for the parameterisation of vertical mixing in the model. Boundary layer schemes include the Yonsei State University (YSU) scheme (Hong et al., 2006), where the vertical diffusion coefficient is calculated according to local temperature gradients, and the Mellor Yamada Janjic (MYJ) scheme, which uses prognostic equations for the Reynolds stress and turbulent heat flux terms, and approximates the higher order terms (Mellor and Yamada, 1982). Although several other planetary boundary layer schemes are available, only the YSU and MYJ schemes were tested here, labelled as Y and M in table 9.2 respectively. The vertical velocity fields comparing the modelled cells for the two schemes are shown in figure 9.9.

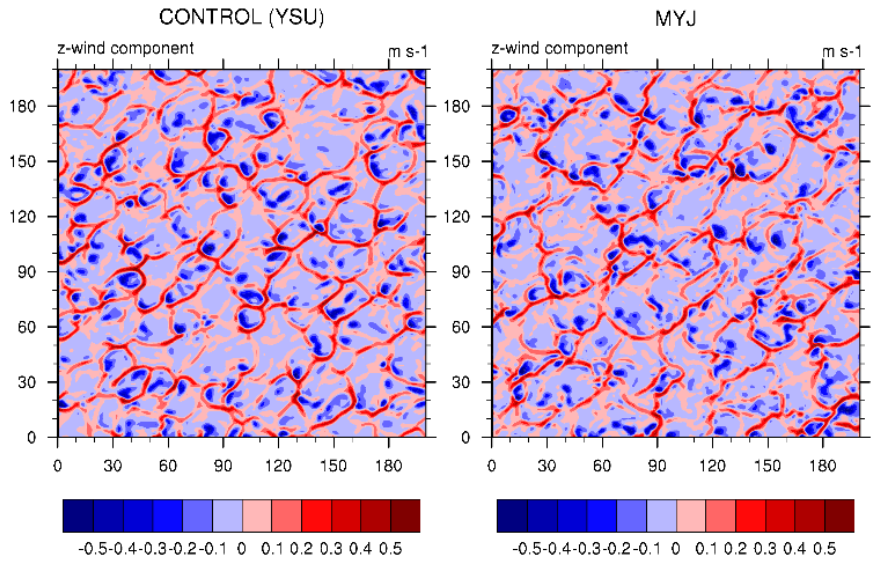


Figure 9.9: Experiments 1 (control) and 9 respectively. Vertical velocity (third model level) for YSU PBL scheme (left) and MYJ PBL scheme (right)

The simulations with the YSU and MYJ PBL parametrizations showed some difference in the structure of the cells and in the amount of high frequency noise, but produced quantitatively similar results. As with the case of the positive definite advection of moisture, there was no way to determine which of these simulations was more realistic or 'correct'. In fact, the primary role of the boundary layer scheme in this case is to dissipate energy from the cells, and to deal with the mixing and entrainment at the top of the boundary layer in a suitable way. It is possible that the simulated OCC are not as sensitive to the boundary layer scheme as some other mesoscale effects, such as the shape of the wind profiles as presented in (Draxl et al., 2010), where the results also suggest that an important difference between the boundary layer schemes in WRF is in the simulation of the transition between day-time unstable and night-time stable conditions.

9.3.4 Positive definite advection of moisture and TKE

Numerical integration schemes in the model can lead to unphysical values of positive definite fields such as moisture or turbulent kinetic energy. For example, Hahn and Mass (2009) demonstrated improved prediction of precipitation for a storm case study when positive definite advection of moisture was used. Experiments 1 (control) and 6 compare the development of the cellular convection patterns with and without positive definite advection of moisture, respectively.

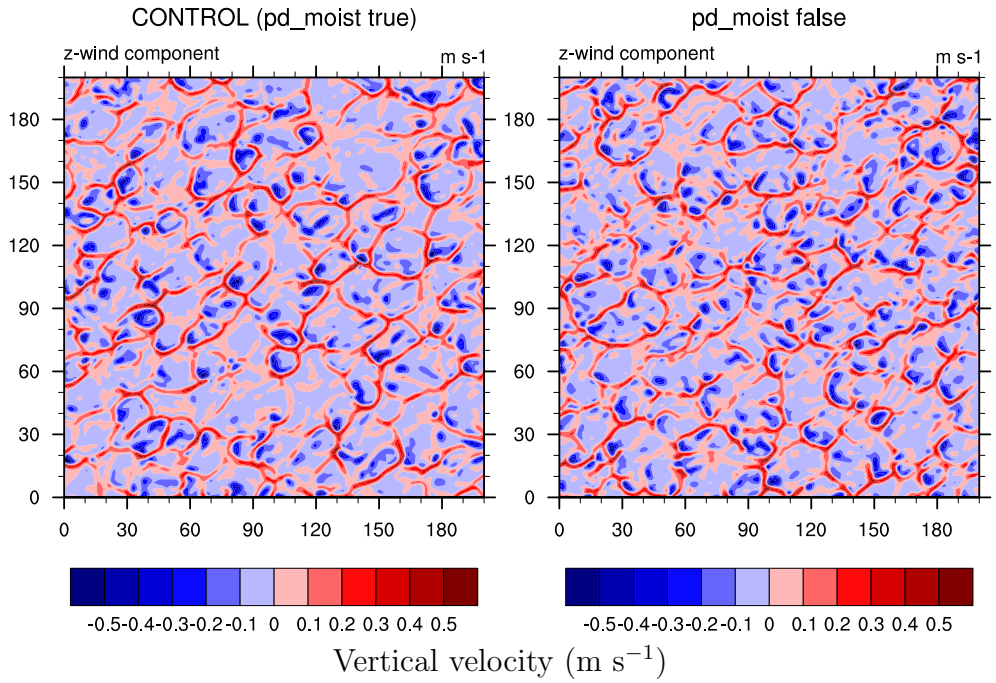


Figure 9.10: Experiments 1 (control) and 6 respectively. Vertical velocity (third model level) with positive definite advection of moisture (left) and without positive definite advection of moisture (right)

The experiments with and without positive advection of moisture showed that there were slight differences in the two simulations, although the difference was difficult to quantify objectively, and showed up only sightly on their kinetic energy spectra. The vertical velocity for the third model level at the end of the 6th simulation hour is shown in figure 9.10. There is some suggestion of a more open, better defined cell structure for the control case with positive definite advection of moisture. This is consistent with the discussion in Wang et al. (2010), where it is suggested that the positive definite advection of moisture is expected to provide some smoothing. It was therefore decided to keep the use of positive definite advection of moisture. Further investigation into the exact role of positive definite advection of moisture would be interesting.

9.3.5 Coriolis force

Whether or not to include the Coriolis force is obviously not a design choice in a real weather model. However, in the idealised model it was not obvious whether to use it or not. Running the model without the Coriolis force neglects its possible influence on cell development, although as will be shown in Chapter 10, it does not turn out to play an important role. On the other hand, including the Coriolis force in the model either means that the wind direction will rotate in a clockwise direction, or that a large scale pressure gradient must be imposed on the domain. A third option available in WRF is to apply the Coriolis force to perturbations in wind speed. These three options were tested here - inclusion of the full Coriolis force (labelled 1 in table 9.2), no Coriolis force (0) and the Coriolis force applied only to the perturbation in wind speed (P).

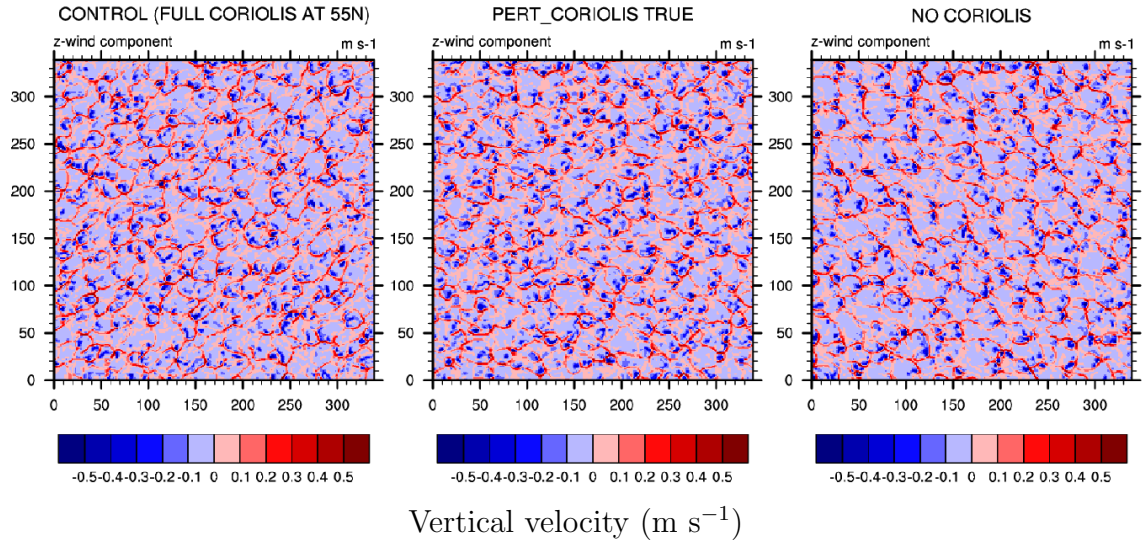


Figure 9.11: Experiments 1 (control), 7 and 8. Vertical velocity (third model level) for full Coriolis (left), perturbation Coriolis (centre) and no Coriolis (right)

When the Coriolis force was turned off, the cells and the mean wind did not rotate as the simulation progressed, and there was qualitatively a minor difference in cell structure. In the case of the perturbation Coriolis option, the cells took on a less organised appearance with more high frequency noise. The turning of the cells with time did not actually present any significant problem, since the domain and surface conditions were homogeneous. Therefore, the full Coriolis force option was used in the revised model setup.

9.4 Physical conditions influencing open cellular convection in WRF

As discussed in chapter 6, several observational studies have investigated one or more dimensions of the parameter space of atmospheric conditions in which mesoscale phenomena such as open and closed cellular convection tend to occur.

For example, the boundary layer height has been studied for its relationship to cell diameter, since the height of the cells is close to the convection height. Assuming a constant cell aspect ratio, this should yield a simple rule for the cell diameter. Bakan and Schwarz (1992) collected data relating cell diameter to convection height using 4 years of satellite imagery over the north-eastern Atlantic ocean together with surface observations and radiosondes from 'Weathership M' (66°N 2°E). They demonstrated that there are no straightforward or simple relations between convection height and cell diameter, and the weak trend between the two parameters shows a great deal of scatter.

The strength of the boundary layer inversion is also an interesting parameter. If there is a strong capping inversion, then momentum and moisture transport out of the boundary layer will be limited. On the other hand, if the boundary layer is only capped by a weak inversion, or by a gradual increase in potential temperature such as in figure 9.1, then the behaviour of the cell flow at the top of the boundary layer is much more complicated.

Another obvious parameter that has been targeted for its influence on cell development is the temperature difference between the air near the surface and the sea water. In the laboratory case of flow between two differentially heated plates, the sea surface temperature is the equivalent of the heating applied to the lower plate. Results from Bakan and Schwarz (1992) show that OCC has been observed for a wide range of air-sea temperature differences between about 2°C and 10°C, and for a wide range of wind speeds between 2 and 25 m s^{-1}.

The purpose of this section is to study the impact of boundary layer height, sea surface temperature and capping inversion on the modelled OCC in WRF. These dependencies map only a small subset of the parameter space of atmospheric conditions for OCC. This is partly because the various parameters do not operate in isolation. The full description of the modelled atmospheric conditions in which OCC occurs is multi-dimensional, and outside the scope of this work.

It is questionable to what extent the results reported here can be transferred directly to the real atmosphere, since there are many aspects of OCC that are not captured in a mesoscale model. These include entrainment, precipitation and turbulent transport associated with the individual cumulus clouds that constitute the cell walls, which are only taken into account in an aggregated sense. To take these factors into account explicitly and separately, an LES model or further observational evidence is required.

For example, de Roode et al. (2004) used an LES model to map the parameter space of the length scale of mesoscale fluctuations as a function of the shape of the vertical heat flux profile and the boundary layer height for a clear convective boundary layer and a stratocumulus topped boundary layer. They found that the length scales of humidity and temperature tended to broaden throughout their simulations, while the length scales of vertical velocity did not. While their study relates to closed cells, in contrast to the open cells considered here, their results show the complexity that exists in modelling mesoscale convection if turbulent fluctuations are taken into account. In the mesoscale simulations presented here, the vertical velocity has the same length scale as all the other fields, because all fields are dominated by the simplified cellular convection. In an LES simulation, the vertical velocity would not show simple, homogeneous regions of upwards and downwards vertical velocity, but would be a complex network of updrafts and downdrafts of many different length scales.

Despite the simplicity of the mesoscale model, the results can be interpreted as a) qualitative confirmation that the response of the model is reasonable b) a guide to what *may* also occur in the real atmosphere c) a demonstration of the sensitivity of the modelled OCC to small changes in input parameters and d) evidence of the potential practical skill of mesoscale models in predicting mesoscale wind fluctuations. Furthermore, as will be demonstrated in chapter 10, the sources and sinks of energy within the modelled OCC can reveal important information about the diffusion in the model in general.

9.4.1 Boundary layer height

Experiments with the boundary layer height were conducted by modifying the profiles of potential temperature and water vapour mixing ratio that were presented in figure 9.1 for the control simulation. For the purposes of these experiments, the 'boundary layer height' was defined as the top of the mixed layer - that is, the point where the potential temperature

began to increase.

Four modifications to the boundary layer height were made by lowering or raising the height of the height of the mixed layer. Above this layer, a profile of similar shape to the original profile was adopted. The water vapour mixing ratio was modified by altering the dewpoint temperature profile on an aerological diagram, and transforming the modified profile back to water vapour mixing ratio. The shape of the dew point temperature profile was maintained as closely as possible to the original profile.

Modifying the profiles of potential temperature and water vapour mixing ratio as described above highlights the difficulty in testing the impact of a single parameter such as the boundary layer height. The boundary layer height does not stand isolated from other aspects of the temperature and water vapour mixing ratio profiles, and the modifications also caused the shape of the profile to change as well as the surface water vapour. With regard to the surface water vapour, this was not seen as a significant disadvantage, since there is an ample availability of moisture over the sea anyway, and the profiles quickly adjusted themselves. The four modified profiles are shown in figure 9.12, relative to the original profile of the control simulation.

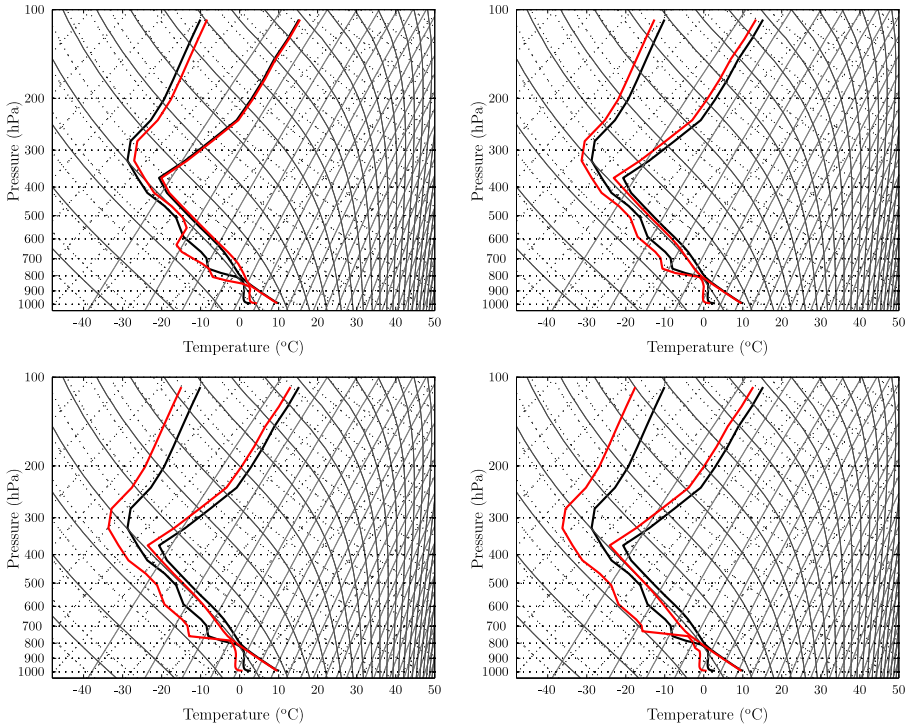


Figure 9.12: Original (black) and modified (red) profiles of temperature and dewpoint temperature.

The vertical velocity fields for the third model level are shown for the end of the 6th simulation hour for the five different temperature and water vapour mixing ratio profiles (the control, plus the four profiles shown in figure 9.12). Since the profiles evolved as the experiments progressed, the boundary layer height at the end of the 6th simulation hour was not exactly the same as the boundary layer height at the beginning of the simulations. The five plots are labelled with the instantaneous boundary layer height, averaged over the whole domain, at the end of the 6th simulation hour.

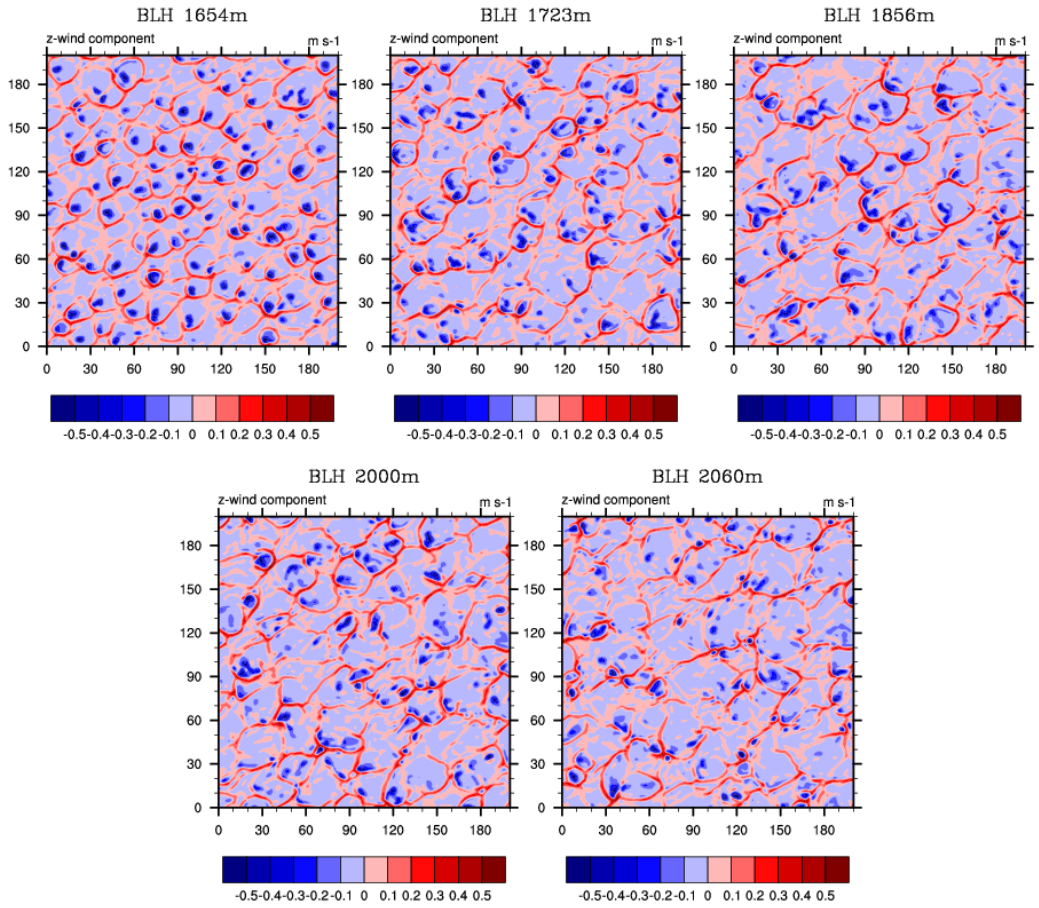


Figure 9.13: Vertical velocity (3rd model level) for 5 different boundary layer heights at the end of the 6th simulation hour. The boundary layer height labelled on each plot refers to the domain-averaged boundary layer height at the end of the 6th simulation hour.

The fields of vertical velocity qualitatively show that the cell diameter increases with boundary layer height. This result is also shown in the spectra of the five fields shown in figure 9.14. In this case, the spectra are shown on a linear axes so that the locations of the peaks are clearer. The spectra show that spectral peak occurs at higher wavenumbers of the lower boundary layer heights, but that the amplitude of the peak does not follow a distinctive pattern. That is, for these experiments, the frequency of the wind fluctuations increases as the boundary layer height increases.

One difficulty with the analysis was that it was impossible to be objective about what stage of the cells' lifecycle was occurring in the different experiments, despite the fact that all the analysis took place after 6 hours of simulation. The boundary layer height and cell diameter for the five cases are shown as a scatter plot in figure 9.15, where the cell diameters were estimated from the peak in the kinetic energy spectra, and the boundary layer height was domain averaged value from the WRF simulation for the end of the 6th simulation hour. No attempt to fit a curve to these points has been made, because it would be incorrect to suggest that there is a simple relationship between two parameters. The points also only cover a small subset of the possible cell diameters and convection heights in comparison to

the extensive data set covered by Bakan and Schwarz (1992), although his results come from the north east Atlantic ocean and are not necessarily representative of the conditions over the North Sea.

As the boundary layer height increases, so does the potential for stronger updrafts, some of which penetrate the top of the boundary layer. It would therefore be expected that the strength of the wind fluctuations should increase with increasing boundary layer height. This was tested by calculating the variance in horizontal wind speed for the third model level at the end of the 6th simulation hour. In figure 9.16, the boundary layer height is plotted against variance in horizontal wind speed, where it seen that the variance generally increases with increasing boundary layer height. However, the relationship is not clearly defined, and the variance actually decreases for the highest boundary layer height of 2060 m.

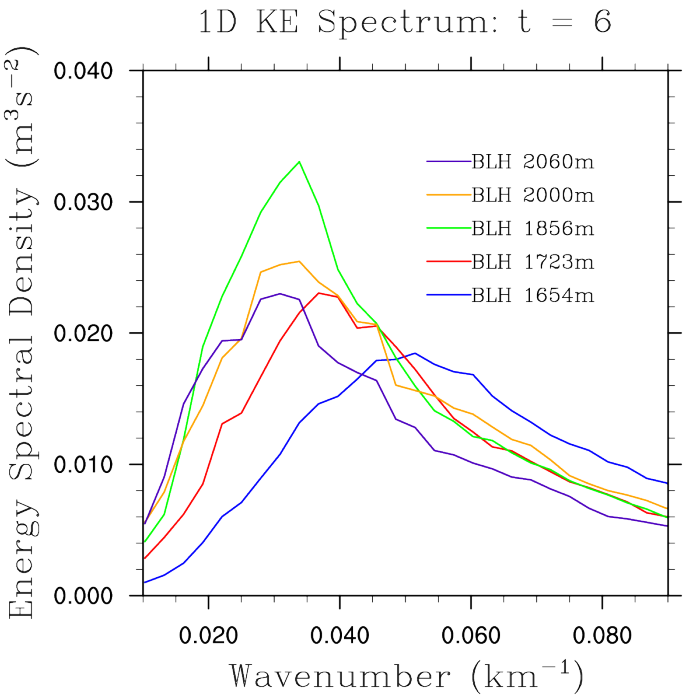


Figure 9.14: Kinetic energy spectra of the five boundary layer height experiments, for the third model level at the end of the 6th simulation hour.

The increase in cell diameter with boundary layer height was not true at all simulation times. For the first few hours, cells that were clearly dependent on the grid spacing developed with only minor differences in intensity between the 5 experiments. Later in the simulation (from hour 8 onwards), the cells lost their structure and the relationship between cell size and boundary layer height almost disappeared.

9.4.2 Sea surface temperature

The horizontally homogeneous sea surface temperature was added to the WRF namelist for the control simulations. Simulations for five different sea surface temperatures were conducted: 282.0 K, 283.9 K, 285.9 K (control), 287.9 K and 290 K. Since the air temperature at the first model level (about 14 m) was 282.6 at the beginning of the simulations, these

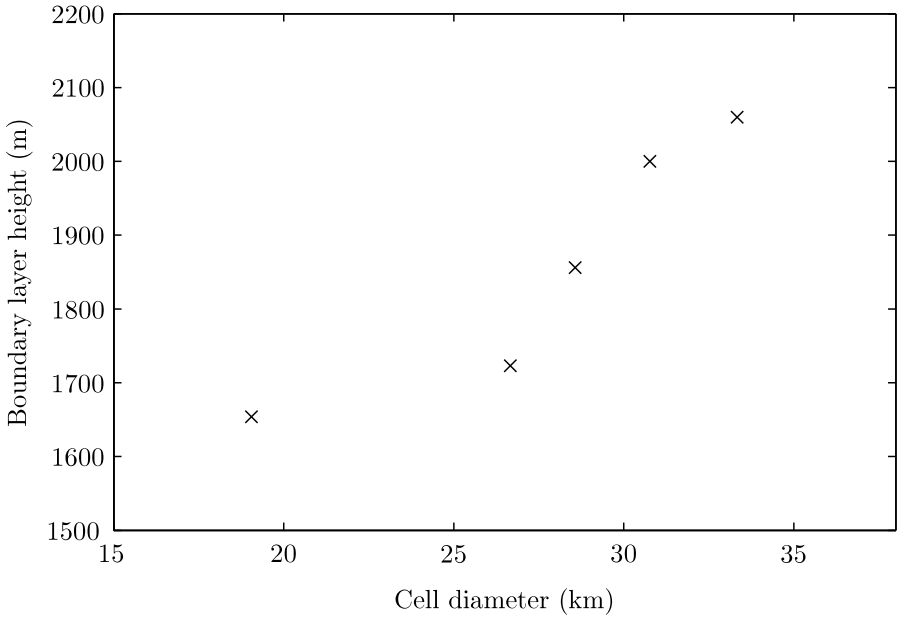


Figure 9.15: Scatter plot of boundary layer height against cell diameter. The cell diameter is determined from the peak of the kinetic energy spectra, and the boundary layer height is from the WRF simulation at the end of the 6th simulation hour.

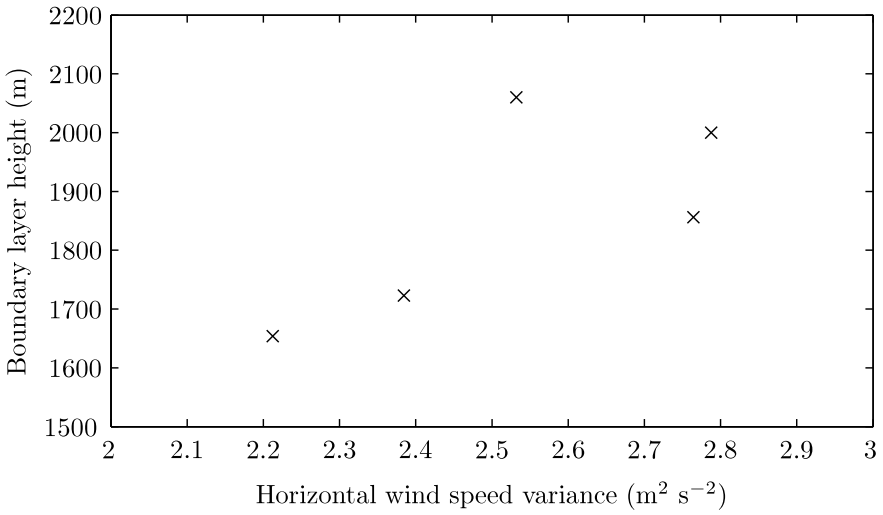


Figure 9.16: Scatter plot of boundary layer height against variance in horizontal wind speed for the third model level at the end of the 6th simulation hour. The boundary layer height is from the WRF simulation at the end of the 6th simulation hour.

sea surface temperatures correspond to air-sea temperature differences of -0.6 K, 1.3 K, 3.3 K (control), 5.3 K and 7.4 K respectively.

The vertical velocity on the 3rd model level for the five different sea surface temperature experiments is shown in figure 9.17. The plots show that the convection fails to develop altogether for the first experiment, where the air-sea temperature difference was -0.6 K. For the other experiments, which had air-sea temperature differences of 1.3 K, 3.3 K, 5.3 K and 7.4 K respectively, the cells are similar but increase in intensity as the air-sea temperature difference increases. For the last experiment, strong and possibly physically unrealistic up-

drafts extended well into the troposphere at many points across the domain. Despite this behaviour of the model, Bakan and Schwarz (1992) did observe cases with an air-sea temperature difference as high as 11 K. This raises the interesting question of how the model would react in a real case where there were very large air-sea temperature differences.

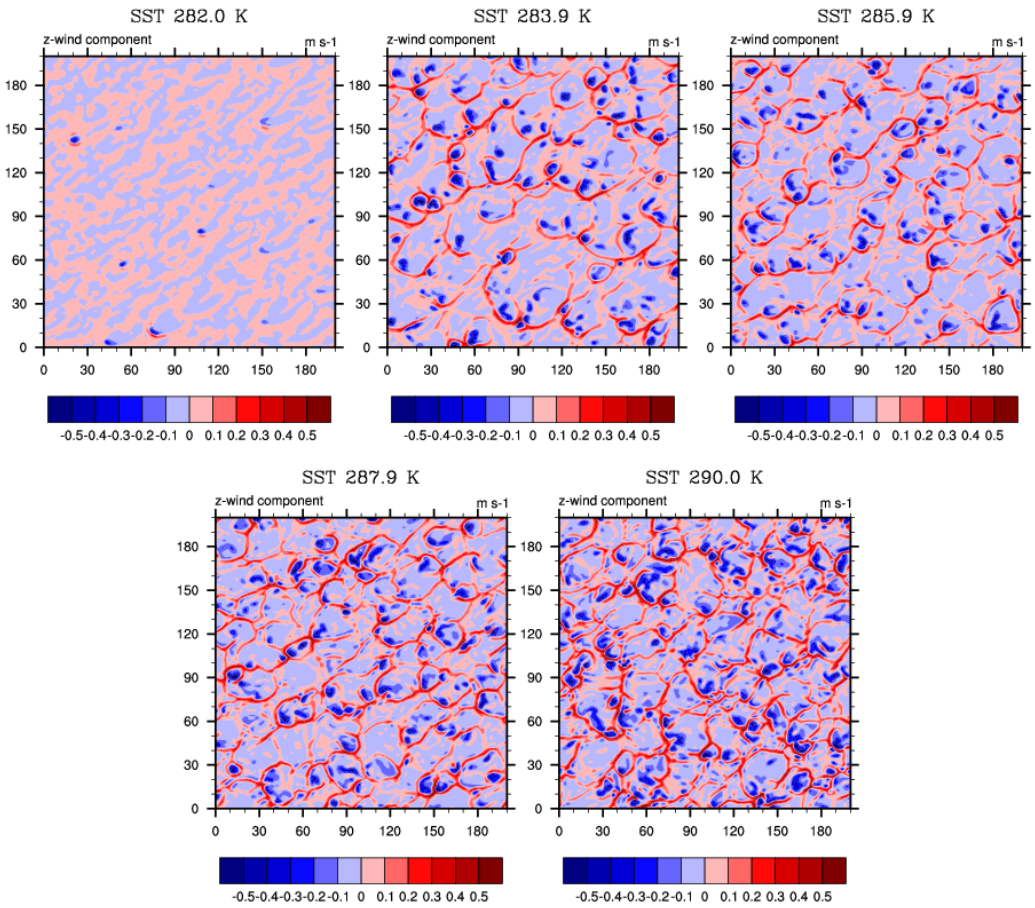


Figure 9.17: Vertical velocity (3rd model level) for five different sea surface temperatures at the end of the 6th simulation hour.

9.4.3 Capping inversion

An artificial inversion was added just above the height of the boundary layer. As for the experiments with boundary layer height, the original profile was modified on a log-P skew-T diagram, so that it was obvious that the dewpoint temperature and potential temperature were consistent with one another. Below the height of the inversion, both profiles were left untouched, as shown in figure 9.18. Above the inversion, the atmosphere is still unstable, but the inversion adds enough convective inhibition to prevent the boundary layer convection from penetrating the top of the boundary layer. Although the method for adding the inversion was purely by-eye and not physically based, the shape of the profiles above the inversion was not important, since the flow was dominated by the development of cells below the inversion.

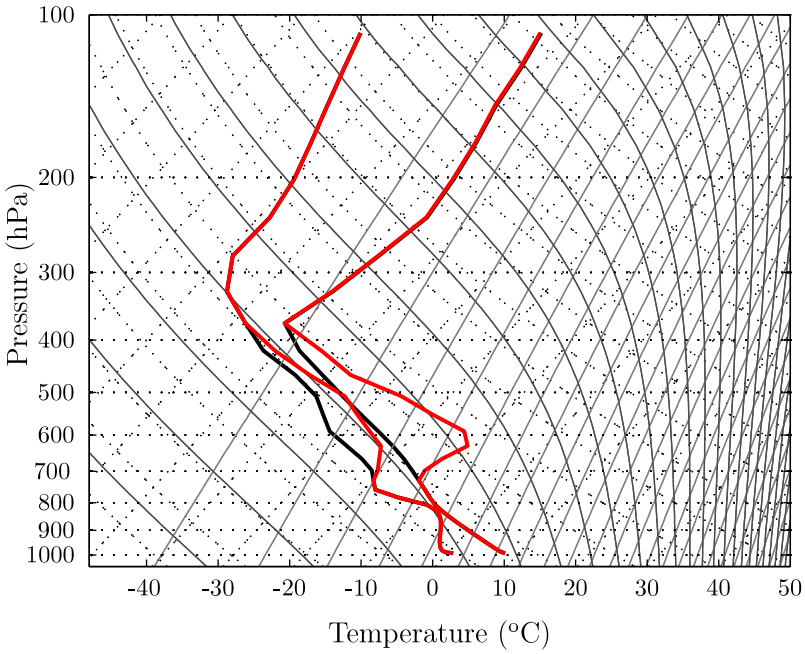


Figure 9.18: Original (black) and modified (red) profiles of temperature and dew point temperature.

The vertical velocity for the 3rd model level for the control case and the case with the capping inversion are shown in figure 9.19. The results show slightly smaller cells for the inversion case, and slightly weaker upwards vertical velocities at the cell walls. This is because the potential for large updrafts overshooting the top of the boundary layer height has been removed. The difference between the simulation with and without the capping inversion will be discussed in further detail in chapter 10, where the cell-scale kinetic energy budgets for the two cases reveal that there is a continued contribution to the mesoscale kinetic energy above the boundary layer height when no capping inversion is present.

9.5 Discussion and Conclusions

In this chapter, the setup and systematic testing of an idealised WRF model was described, and the results of sensitivity tests of the model were presented. In the simulations, OCC developed that was similar to that in the simulated real weather cases that were presented in chapter 8. Cells had a diameter of around 20 km, with large upwards vertical velocity in the cell walls and downwards vertical velocity in the cell centres. Due to the low resolution of the model, it was not expected that the individual cumulus clouds that make up the cell walls would be simulated. This shortcoming did not appear to prevent the development of the overall structure of the cells.

Systematic testing of the idealised model guided a revised setup, which was used as the control case for a set of sensitivity experiments of the modelled OCC to the environmental forcing. A series of experiments were conducted that tested the impact of boundary layer height, capping inversion and sea surface temperature on the development of OCC in WRF.

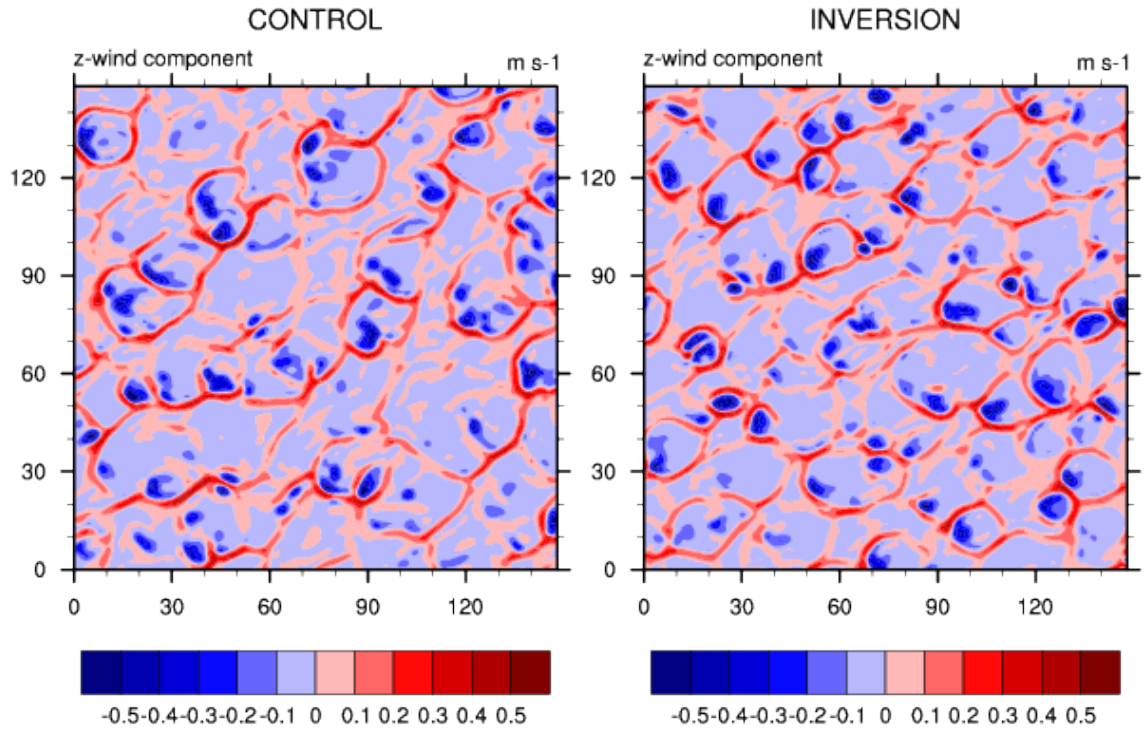


Figure 9.19: Vertical velocity (3rd model level) for the cases without (left) and with (right) the capping inversion, at the end of the 6th simulation hour.

The results cannot be mapped directly to the equivalent impacts in the real atmosphere because a mesoscale model does not contain high frequency information and only approximates many of the physical processes in the boundary layer. However, the results do provide insight into the trends in these impacts, and also highlight the sensitivity of the model to parameters such as sea surface temperature and boundary layer height. Interestingly, these two factors are often only coarsely defined in the initial conditions of the model. In the case of the boundary layer height, it was found that changes of only a couple of hundred metres made a significant difference to the spatial scale of the modelled structures, and therefore to the frequency of the associated wind fluctuations. Such differences in boundary layer height could easily come about from errors in initial conditions, or simply as a result of interpolation from the coarse resolution large scale model from which the initial conditions are taken. Therefore, these results provide interesting evidence of the importance of boundary layer height in mesoscale model initialisation, and possibly point towards the benefit of data assimilation of well resolved boundary layer observations.

Similarly, the sea surface temperature is often obtained from rather coarse data and treated as a constant throughout a short mesoscale model simulation. The results presented here demonstrate the importance of correct and detailed sea surface temperature information, particularly in areas where the shallow water depth allows a well developed diurnal cycle in water temperature, such as around areas of the North Sea coast.

Experiments equivalent to those presented here but conducted using an LES model would be a fascinating extension to the work. As pointed out by de Roode et al. (2004), the serious limitation to this approach is the LES domain must be large enough that the mesoscale

structures can be represented such that they are not influenced by the domain size itself. In their LES simulations of clear convective boundary layers and stratocumulus topped boundary layers (which had a domain size of 25.6 km and a horizontal grid spacing of 100 m), de Roode et al. (2004) found that the length scales of potential temperature and moisture increased throughout their 8 hour simulation, but that the vertical velocity and virtual potential temperature maintained length scales in the order of the convection height. In the mesoscale case presented here, the length scale of the boundary layer height (around 2000 m) is not even resolved in the model, and the vertical velocity has a spectral peak relating to the cell diameter. It is not clear whether the expansion and ultimate decay of the cells in the mesoscale model is related to the cell broadening discussed by de Roode et al. (2004), or whether it is simply an artifact of the experimental design.

The results presented here show only a small subset of the multi-dimensional set of atmospheric conditions that could be tested for their influence on the development of OCC. One of the main difficulties that was encountered during the experiments was the problem of the simulation environment modifying itself during the course of the experiments. This meant that it was difficult to effectively control the boundary layer height that was being tested, since the boundary layer height of the profile used to initialise the model also changed during the simulation. The eventual decay of the cells in the simplified WRF setup does not parallel that in the real world, and may not even resemble what happens in the full WRF model, because the idealised experiment is a closed system without any supply of new, unstable air, as would be the case in both the real atmosphere and in the full WRF model. However, although the boundary layer heights did evolve during the experiments, they maintained their order from lowest to highest and still allowed a comparison of the cell structure at difference boundary layer heights. Solutions to this problem could involved specifying the heat flux in the environment directly, or working backwards so that the model was initialised in such a way that the desired conditions occurred after some fixed number of hours.

Chapter 10

Energetics of the modelled open cellular convection

10.1 Introduction

In chapter 9, an idealised simulation of open cellular convection (OCC) over water was presented, and it was demonstrated that the modelled convection is highly sensitive to input parameters such as boundary layer height and sea surface temperature. The purpose of this chapter is to explore the physical forcing for the OCC through the calculation of the cell scale kinetic energy budget. This work has important implications not only for understanding the physical processes behind of the modelled OCC, but for observing the way that energy is dissipated in the model itself. The analysis could be extended to give important modelling insights, such as the performance of different boundary layer schemes near the top of the boundary layer. Furthermore, OCC is an important mechanism for transport of energy between the surface the upper atmosphere, and the comparison of the results presented here with observations could reveal interesting aspects of this energy transfer, both in relation to high resolution mesoscale models which explicitly simulate the cells, and in large scale and climate models which do not explicitly model the OCC itself.

The precise identification and analysis of the physical processes that are simulated in a mesoscale model is also an important step for model development and verification, since the impact of such processes on surface wind fields will be masked in traditional time series and spatial verification methods. Not only will the wind fluctuations be spectrally blended with other mesoscale and topographically forced features, but the fluctuations will generally not have the correct phase and will therefore be heavily penalised when using verifications such as the root mean square error.

LeMone (1973) and LeMone (1976) calculated a roll-scale kinetic energy budget for convective rolls based on observational data by decomposing the flow into mean, roll-scale and turbulent components. Brümmer (1985) later reproduced her budget for observed cases of convective rolls over the North Sea. In both cases, it was found that production of roll-scale kinetic energy by shear and buoyancy were the most important terms in the maintenance of the

rolls, although of three cases studied by Brümmer (1985), two appeared to be shear-driven and one appeared to be buoyancy driven.

Cieszelski (1998) followed similar reasoning to calculate buoyancy and shear terms in the cell scale kinetic energy budget for OCC, although the budget is more complicated in this case because the problem is 3 dimensional, in contrast to the 2 dimension case of convective rolls. He used aircraft observations collected over the North Sea during the KonTur experimental campaign for a case where OCC of diameter 2–3 km could be identified from satellite and aircraft photographs. The cells that he studied were therefore much smaller than those studied here, but his basic methodology could still be adapted to the simulated OCC in WRF.

LeMone (1973, 1976), Cieszelski (1998) and Brümmer (1985) all used observations to calculate the roll-scale or cell scale kinetic energy budget, which meant that they filtered the turbulent fluctuations from the data with a filter for which they knew the spectral attenuation. In our case, there are no turbulent fluctuations in the mesoscale model, and we assume that the cell scale fluctuations are in fact that smallest scale that is represented in the model. The disadvantage of this assumption is that we rely on the implicit and explicit filtering in the model, for which the attenuation frequency is neither exactly nor cleanly defined.

An important difference between the OCC simulated in a mesoscale model and that in the real atmosphere is that in the real atmosphere, the cell walls are made up of rings of individual cumulus clouds, while the simulated cell walls are relatively homogeneous areas of upwards vertical velocity. This means that small scale effects such as entrainment and modulation of the boundary layer height at the top of individual cumulus updrafts are absent from the mesoscale simulation, as are rain showers associated with individual clouds. These issues are particularly important for the calculation of the energy budget, because although the mixing due to turbulence is parameterised in the model, and the energy dissipation arising from turbulence still needs to be taken into account in the budget.

In this chapter, the cell scale kinetic energy budget is calculated for two of the simulations that have already been presented in chapter 9. The two cases are the control case and the case with the strong inversion added just above the boundary layer height. A further discussion of the cases can be found in section 9.4.3. The only differences between the cases presented in figure 9.19 and the ones used here are that the flux limiter for the 6th order diffusion is not used and that a 4th order vertical advection scheme and a 6th order horizontal advection scheme are adopted, as discussed further in section 10.3. This is because the impact of the flux limiter on the energy balance of the model is difficult to quantify, and because the even ordered advection schemes simplify the calculation of the total diffusion in the model, as discussed later in this chapter.

10.2 Theoretical background to the cell scale kinetic energy budget

The expression for the mesoscale kinetic energy budget, which closely follows that given in Cieszelski (1998), is derived from the momentum and continuity equations by decomposing the wind vector, $\mathbf{V}$, the virtual potential temperature, θ_v and the pressure, p into three

components consisting of the mean flow, the mesoscale fluctuations associated with the OCC, and the turbulent fluctuations,

$$\mathbf{A} = \mathbf{A}_\infty + \mathbf{A}_l + \mathbf{A}', \tag{10.1}$$

where $\mathbf{A}_\infty$ refers to the mean flow, $\mathbf{A}_l$ refers to the cell scale fluctuations and $\mathbf{A}'$ are the turbulent fluctuations.

This decomposition was applied to the Navier-Stokes equations (equation 10.2) for incompressible flow.

$$\frac{\partial \mathbf{V}}{\partial t} = -(\mathbf{V}\cdot\nabla)\mathbf{V} - \frac{1}{\rho}\nabla p - 2\Omega_e \times \mathbf{V} + g\frac{\theta_v}{\theta_{v,0}}\mathbf{k} \tag{10.2}$$

In the idealised simulation of OCC, there are no scales larger than the cell scale flow, so the mean flow is constant by definition in the horizontal directions. This means that $\frac{\partial U_\infty}{\partial x} = \frac{\partial U_\infty}{\partial y} = \frac{\partial V_\infty}{\partial x} = \frac{\partial V_\infty}{\partial y} = 0$. We also assume that there is no large scale vertical motion, although in the actual simulation, W_∞ is of order 10^{-5} m s^{-1}. The mean flow is assumed to be constant with time, despite the fact that it is turning slowly due to the Coriolis force, so $\frac{\partial U_\infty}{\partial t} = \frac{\partial V_\infty}{\partial t} = \frac{\partial W_\infty}{\partial t} = 0$. The continuity equation is assumed to be valid for each of the three scales, so that

$$\nabla\cdot\mathbf{V}_\infty \approx \nabla\cdot\mathbf{V}_l \approx \nabla\cdot\mathbf{V}' \approx 0. \tag{10.3}$$

Substituting the decomposition in equation 10.1 into equation 10.2 and multiplying by $\mathbf{V}_l$, an equation for the mesoscale component of the kinetic energy is obtained:

$$\frac{\partial E_l}{\partial t} = \mathbf{V}_l\frac{\partial \mathbf{V}_l}{\partial t}. \tag{10.4}$$

Filtering the budget by averaging over the whole domain means that the usual spatial averaging rules for turbulent fluctuations apply. Furthermore, we assume that turbulent fluctuations are uncorrelated with cell scale fluctuations, so that the relevant set of averaging rules are

$$\begin{gathered}\widehat{a'} = 0, \quad \widehat{a_l} = 0, \quad \widehat{a'b_l} = 0, \quad \widehat{a'b_\infty} = 0, \\ \frac{\partial}{\partial x_i}\widehat{a'} = 0, \quad \frac{\partial}{\partial x_i}\widehat{a'b_l} = 0 \quad \text{and} \quad \frac{\partial}{\partial x_i}\widehat{a'b_\infty} = 0,\end{gathered} \tag{10.5}$$

where a and b are any combination of the three velocity components, pressure or virtual potential temperature, and the $\widehat{\ }$ denotes filtering by averaging over a sufficiently large number of cells. In this case, the filtering is achieved by horizontal averaging over the whole domain, which is assumed to be dominated only by the cell scale flow.

Expanding equations 10.2 and 10.4, together with the assumptions discussed above, gives a budget for time rate of change rate of the mesoscale component of the domain averaged kinetic energy. In the case of the idealised WRF simulations, this component is uniquely attributable to the OCC, since we assume that there are neither other mesoscale features or sub-cell scale fluctuations. The full derivation of the equation is given in appendix B. The domain averaged mesoscale kinetic energy budget is similar to that for convective rolls in LeMone (1973, 1976) and for OCC in Cieszelski (1998), although Cieszelski's formulation of the pressure-wind balance term is different to the one given here.

$$\frac{\partial \hat{E}_l}{\partial t} = -U_\infty \frac{\partial \hat{E}_l}{\partial x} - V_\infty \frac{\partial \hat{E}_l}{\partial y} \underbrace{- \widehat{u_l w_l} \frac{dU_\infty}{dz} - \widehat{v_l w_l} \frac{dV_\infty}{dz}}_{\text{shear}} + \underbrace{\frac{g}{\theta_{v,\infty}} \widehat{\theta_{v,l} w_l}}_{\text{buoyancy}}$$

$$\underbrace{- \widehat{u_{j,l}} \frac{\partial \hat{E}_l}{\partial x_j}}_{\text{cell scale advection}} \quad - \quad \underbrace{\frac{1}{\rho} \frac{\partial}{\partial x_j} \widehat{p_l u_{j,l}}}_{\text{pressure-wind balance}} \quad \underbrace{- u_{k,l} \frac{\partial}{\partial x_j} (\widehat{u'_k u'_j})_l}_{\text{interscale transfer}} . \tag{10.6}$$

The first three terms in equation 10.6 are the Lagrangian rate of change of the cell scale kinetic energy. The shear terms are the generation of cell scale kinetic energy due to shear in the mean flow, and the buoyancy term is the generation of cell scale kinetic energy due to buoyancy. The cell scale advection term is the advection of cell scale kinetic energy due to the cell scale flow itself. The pressure-wind balance is a restoring term which acts against cell scale flow that is counter-gradient to the cell scale pressure fluctuations. The final term, the interscale transfer, is the loss or gain of energy to turbulent scales. As discussed by LeMone (1973, 1976) and Cieszelski (1998), this term appears in the turbulent kinetic energy budget with the opposite sign. Although in the real atmosphere energy is lost through molecular dissipation, in a mesoscale model this term accounts for all the energy in the mesoscale model that is ultimately lost through explicit and numerical diffusion. The Coriolis terms cancel one another in the cell-scale energy budget, in contrast to the two dimensional budget of LeMone (1973, 1976), who only calculated the energy of the j and k components of the roll-scale energy budget, and was therefore left with two Coriolis terms in her equation.

10.3 Calculation of the simulated mesoscale kinetic energy budget

The domain averaged mesoscale kinetic energy budget was calculated after five hours of simulation. All fields were first averaged to levels of constant height. The terms involving only u_i, $U_{i\infty}$, θ_{vi} and $\theta_{v\infty}$ were straightforward to calculate, since the cell scale fluctuation terms were derived directly by subtracting the mean flow at each level from the total flow - that is, the buoyancy, shear, cell scale advection and pressure-wind balance terms. $\frac{\partial E}{\partial t}$ was estimated directly from a centered difference, with a discrete time step of one minute.

The final term in the budget, the interscale transfer (IST), was the most involved to calculate. Written out in full, the interscale transfer term (IST) in the mesoscale kinetic energy balance is

$$IST = \begin{bmatrix} \frac{\partial}{\partial x}(\widehat{u'u'}) + \frac{\partial}{\partial y}(\widehat{u'v'}) + \frac{\partial}{\partial z}(\widehat{u'w'}) \\ \frac{\partial}{\partial x}(\widehat{v'u'}) + \frac{\partial}{\partial y}(\widehat{v'v'}) + \frac{\partial}{\partial z}(\widehat{v'w'}) \\ \frac{\partial}{\partial x}(\widehat{w'u'}) + \frac{\partial}{\partial y}(\widehat{w'v'}) + \frac{\partial}{\partial z}(\widehat{w'w'}) \end{bmatrix} \cdot \mathbf{V}_l \tag{10.7}$$

where the nine terms constitute sub-grid scale momentum fluxes. As in all mesoscale models, the sub-grid scale fluxes in WRF are parametrised by the horizontal diffusion, the vertical diffusion from the planetary boundary layer scheme, the optional 6th order explicit diffusion and, depending on the order of the advection scheme, the implicit numerical diffusion. In WRF, the first three of these sources of diffusion are added to the three velocity components as additional tendency terms in the Navier-Stokes equations. All three sources are parameterised as a diffusion constant multiplied by a partial derivative, so that the explicit diffusion contribution to equation 10.7 can be written as

$$IST_{explicit} = \begin{bmatrix} \frac{\partial}{\partial x}(K_m \frac{\partial U}{\partial x}) + \frac{\partial}{\partial x}(K_6 \frac{\partial^5 U}{\partial x^5}) + \frac{\partial}{\partial y}(K_m \frac{\partial U}{\partial y}) + \frac{\partial}{\partial y}(K_6 \frac{\partial^5 U}{\partial y^5}) + \frac{\partial}{\partial z}(K_{pbl} \frac{\partial U}{\partial z}) \\ \frac{\partial}{\partial x}(K_m \frac{\partial V}{\partial x}) + \frac{\partial}{\partial x}(K_6 \frac{\partial^5 V}{\partial x^5}) + \frac{\partial}{\partial y}(K_m \frac{\partial V}{\partial y}) + \frac{\partial}{\partial y}(K_6 \frac{\partial^5 V}{\partial y^5}) + \frac{\partial}{\partial z}(K_{pbl} \frac{\partial V}{\partial z}) \\ \frac{\partial}{\partial x}(K_m \frac{\partial W}{\partial x}) + \frac{\partial}{\partial x}(K_6 \frac{\partial^5 W}{\partial x^5}) + \frac{\partial}{\partial y}(K_m \frac{\partial W}{\partial y}) + \frac{\partial}{\partial y}(K_6 \frac{\partial^5 W}{\partial y^5}) \end{bmatrix} \cdot \mathbf{V}_l \tag{10.8}$$

where K_m is the horizontal diffusion constant, which is calculated according to the Smagorinsky first order closure scheme and K_{pbl} is the vertical diffusion constant, which is calculated from the YSU planetary boundary layer scheme. K_6 is the diffusion constant for the additional sixth order diffusion which is specified in WRF as $\frac{\beta 2^{-6}}{2\Delta t}$ (Knievel et al., 2007).

The tendencies in equation 10.8 were calculated on-line in the model, so that they were the exact tendencies that were applied in the model. This was important, because if they were calculated off-line from the saved model fields at the end of the time step, diffusion would already have been applied and the gradients would have been smoothed compared to the values that they had at the time the diffusion was calculated.

Additional implicit numerical diffusion arises out of the odd-ordered Runge-Kutta advection schemes. The default schemes in WRF are 3rd order upwind vertical advection, and 5th order upwind horizontal advection. As described in Hundsdorfer et al. (1995) and Wicker and Skamarock (2002), the odd-ordered Runge-Kutta advection schemes may be written as the next highest order even-ordered scheme plus a diffusion term, while the even-ordered schemes themselves are non-diffusive. For example, for the advection of a scalar, q, by the U velocity component, the third order upwind Runge-Kutta discretisation is:

$$[U\frac{\partial q}{\partial x}]_i = U\frac{2q_{i-2} - 12q_{i-1} + 6q_i + q_{i+1}}{12dx} \tag{10.9}$$

This can be written as the sum of the fourth order upwind Runge-Kutta discretisation and the centered fourth derivative.

$$[U\frac{\partial q}{\partial x}]_i \approx U(\frac{2q_{i-2} - 8q_{i-1} + 8q_{i+1} - q_{i+2}}{12dx} - \frac{h^3}{12}\frac{q_{i-2} - 4q_{i-1} + 6q_i - 4q_{i+1} + q_{i+2}}{h^2})$$
$$\approx U(\frac{2q_{i-2} - 8q_{i-1} + 8q_{i+1} - q_{i+2}}{12dx} - \frac{dx^3}{12}\frac{\partial^4 q}{\partial x^4}_i) \quad (10.10)$$

Equation 10.10 shows that the implicit diffusion arising from the vertical third order upwind advection scheme is equivalent to fourth order diffusion with a diffusion constant of $W\frac{dz^3}{12}$. Similarly, the implicit diffusion associated with the horizontal fifth order upwind advection scheme is equivalent to sixth order diffusion with diffusion constants of $U\frac{dx^5}{60}$ and $V\frac{dy^5}{60}$ for the two horizontal components. To avoid the problem of explicitly quantifying the numerical diffusion, which could be achieved by adding an on-line calculation to the model at the same point of the time step was the advection calculation itself, even ordered, non-diffusive advection schemes have been used. The schemes were 6th order for horizontal advection and 4th order for vertical advection.

10.4 Results

The domain averaged mesoscale kinetic energy budgets for the two simulations are shown in figure 10.1. Qualitatively, the budgets are similar, but the buoyancy, pressure balance and explicit diffusion terms in the budget have greater magnitudes in the case without the inversion. For the case with the inversion, all terms in the budget become negligibly small immediately above the boundary layer height, while for the case without the inversion there are continued contributions to the growth and decay of mesoscale kinetic energy above top of the boundary layer, which reflect the plumes that overshoot the boundary layer top. Moreover, for the case without the capping inversion, the left and right hand sides of the budget do not balance as closely as they do for the case with the capping inversion. This is most likely due to the exchange of energy with the upper atmosphere.

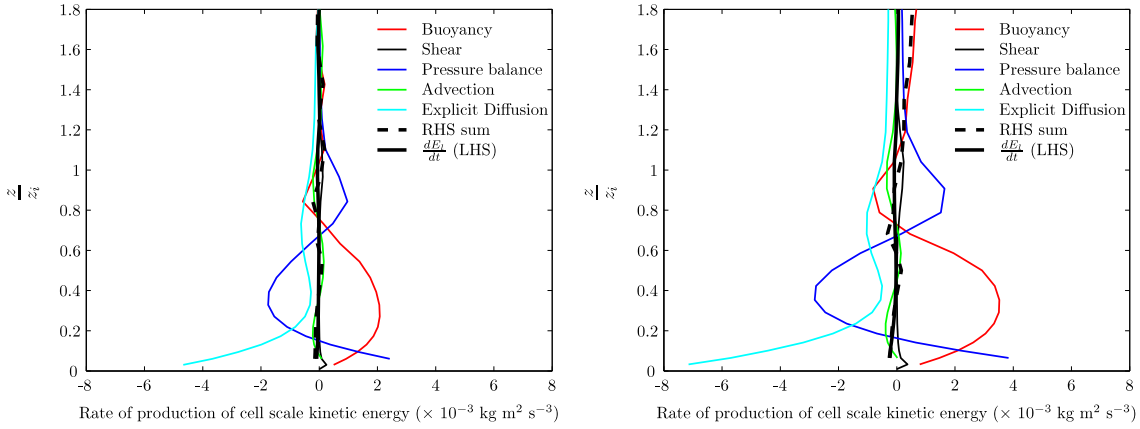


Figure 10.1: The domain averaged mesoscale kinetic energy budget as a function of height for the simulation with strong capping inversion (left) and without strong inversion (right). The vertical axis is height, z, scaled by the boundary layer height, z_i. The shear and advection terms are small compared with the other terms, and are not visible at most heights.

To show the energetics of the cell from a different perspective, each budget term was also averaged over 30 representative cells for a vertical cross section through the cells, perpendicular to the mean wind. This averaging operation takes the place of the whole domain averages shown in figure 10.1, although in this case there are some extra terms of form $A_\infty b_l$ arising from the Coriolis term which do not locally equate to zero inside the cell. These terms vanish when considering the horizontal average over the whole composited cell. The cross sections are for the simulation with the added inversion, since this case shows the flow within the cells in a much clearer way. The chosen cells were the 30 cells with the most clearly defined, closed cell walls. Their mean radius was 8.7 km, with a standard deviation of 1.6 km. Due to this distribution of sizes, the cells were scaled slightly in the horizontal extent, so that they all had the same radius. This will introduce small errors into the analysis, but will not affect the qualitative description of the cell energetics. In figure 10.2, the cross-wind and vertical velocity are shown as vector arrows, and the vertical velocity is also contoured, showing the average cell flow for the 30 cells. In figure 10.3, the six budget terms are shown. In both cases, the vertical axis shows the height, z, scaled by the boundary layer height, z_i, and the horizontal axis shows the average horizontal position, $\hat{y}$, scaled by the average cell diameter, $2\hat{r}$. The strong surface divergence and convergence that is seen in figure 10.2 is associated with the cell-scale pressure fluctuations, which cause the strong contribution to the pressure-wind balance term at the surface (figure 10.3).

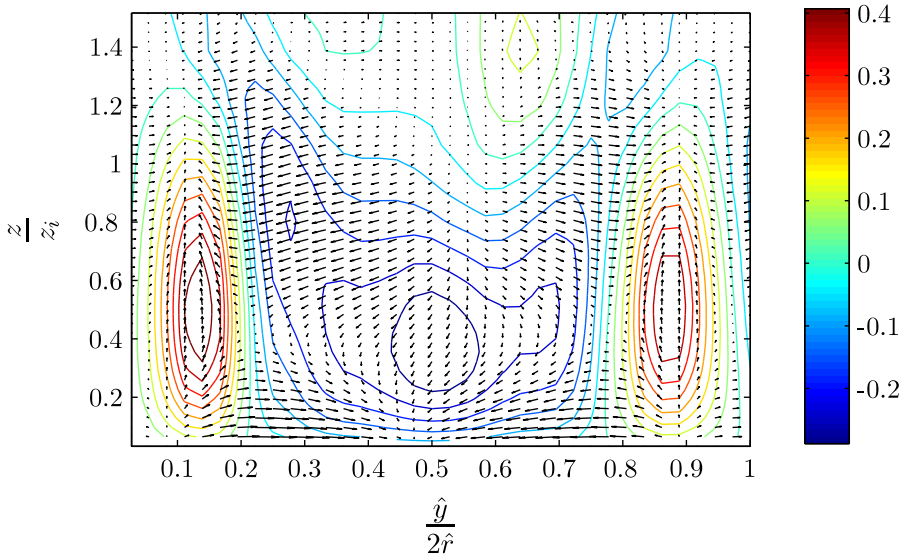


Figure 10.2: Cell scale cross wind and vertical velocity (vectors arrows) and vertical velocity (contours) on a cross section through the cell perpendicular to the mean wind. The flow is averaged over 30 cells. The vertical axis is height, z, scaled by the boundary layer height, z_i, and the horizontal axis is the averaged horizontal position across the cell cross section, scaled by the average cell diameter, $2\hat{r}$.

10.5 Discussion and Conclusions

The cell scale kinetic energy budgets show that the production of kinetic energy by buoyancy is almost balanced by the pressure-balance term in the budget, and that the interscale transfer term is responsible for closing the budget. The shear and advection terms integrate to almost zero over the spatial extent of the cells. These results are somewhat different from

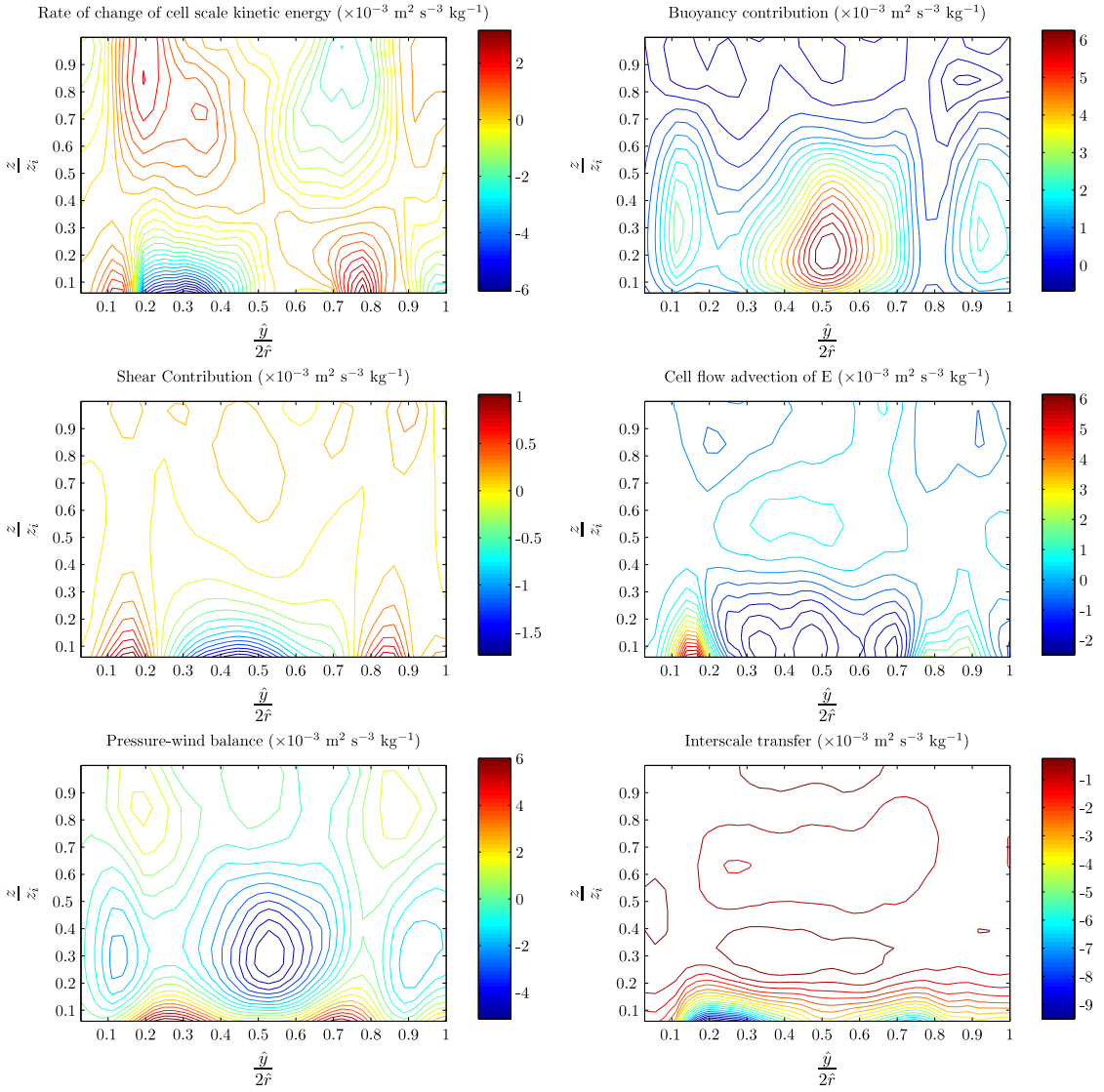


Figure 10.3: Budget terms for the even order advection simulation. All terms are averaged over 30 cells for a cross section through the cell perpendicular to the mean wind. Axis definitions are the same as for figure 10.2. Note that the scales on the contouring are not the same for all terms.

the results of LeMone (1976) and Brümmer (1985), who both found that for convective rolls, the important terms in roll maintenance were shear and buoyancy. Interestingly, neither of them calculated the pressure balance term, with LeMone (1976) stating that these terms, along with the cell flow advection term, *'integrate to zero over the entire roll, provided no energy is transported out the top'*. Similarly, Cieszelski (1998) did not calculate this term.

LeMone (1976), Brümmer (1985) and Cieszelski (1998) all used observational data for their calculations. Here, the use of model data made it possible to calculate every term in the budget. For the first time, a complete and balanced budget of a phenomenon in the family of mesoscale cellular convection has been presented. A major result from the budget is that the pressure balance term is not negligible, but is one of the three important terms that balances the budget. OCC is obviously a different phenomenon from convective rolls, and it would be interesting to repeat the work in this chapter for a modelled case of convective rolls for comparison with the observational results of LeMone (1976) and Brümmer (1985)

Furthermore, the use of model data permitted the calculations of all the budget terms as cross sections through the cells, as well as a cross section of the flow within the cell. The cell cross sections show that the maximum horizontal fluctuations occur just inside the cell walls, which is where the maximum surface divergence must occur. The fluctuation in cross wind velocity is $\pm$ 3 m s^{-1}, which is large enough to explain the fluctuations observed in the horizontal wind speed at the Horns Rev I wind farm during OCC events, particularly considering that there is some smoothing due to averaging over the 30 cells. The cross sections of the budget terms reveal a highly reasonable and plausible story: The buoyancy has local maxima in the cell walls and in the cell centre, where negative fluctuations in both virtual potential temperature and vertical velocity combine to give a positive contribution to the buoyancy term. The shear contribution is concentrated near the surface, where the strongest gradients are found but balances to nearly zero across the cell. The pressure balance term is largest at the surface, where the strong surface divergence results in strong counter gradient flow, and the interscale transfer term is maximum near the surface, which is where the energy should be being dissipated.

Relying on model data is also a serious disadvantage. There is no direct evidence that the budget of modelled OCC is a good model of the budget in the real atmosphere, and the coarse resolution and parametrised boundary layer mean that several important physical processes are simply missing from the calculations. For example, diffusion and energy transfer due to turbulence, surface fluxes and entrainment at the top of the boundary layer are heavily parametrised. All the diffusion in the model is combined into the explicit and implicit diffusion terms that arise out of a combination of the boundary layer scheme, the horizontal diffusion, the additional 6th order horizontal diffusion and, depending on the advection order, the discretised treatment of the advection terms.

A really fascinating study would be to repeat these experiments with a large eddy simulation model, where the larger turbulent eddies would be modelled directly, and the interscale transfer term would actually be a transfer of energy between the meso- and turbulent scales. Even then, an LES model is not the same as the real atmosphere, and extra observational evidence would be ideal. However, given the scale of the cellular convection and the density of measurements that would be required, calculating a fully balanced budget based on observations seems rather prohibitive. Volume scanning remote-sensing techniques such as lidars could be applied to the problem in the future.

The results in this chapter, together with those in chapters 8 and 9, also have implications for the use of mesoscale models for wind energy forecasting. The ability of mesoscale models to simulate OCC means that model forecasts could be usefully applied to the problem of predicting the onset and amplitude of mesoscale wind fluctuations, which is a relevant problem for offshore wind farms. Traditional verification scores such as the root mean square error will not necessary reward such forecasts, despite the fact that they contain physically realistic and potentially useful meteorological information, since the phase of the fluctuations will almost certainly be wrong. Therefore, the use of such forecasts needs to move beyond the paradigm of deterministic time series forecasting, and address instead the second order or spectral information contained in the forecasts. Probabilistic forecasts of such events, for example based on mesoscale ensemble models will further extend the concept of wind variability to include predictability.

Skill in simulating OCC is not isolated from the overall skill of the model. Even if a mesoscale model is successfully able to simulate OCC, it will fail if the large scale forcing does not pro-

vide the correct environment. Features of the large scale flow such as boundary layer height, air-sea temperature difference and boundary layer temperature profile are all important if realistic OCC is to develop with the right length scale, intensity and spatial coverage. This has implications both for the correct downscaling of large scale models, and for the correct initialisation of the boundary layer in mesoscale models through high resolution data assimilation. On the other hand, if a mesoscale model fails in accurately predicting OCC, then this may feedback to large scales in the model, since OCC is clearly an important mechanism for transporting energy from the surface to the top of the boundary layer. These are interesting areas for further study.

Chapter 11

Summary and Conclusions

Mesoscale fluctuations over the sea are an important and interesting aspect of the near-surface wind. They have serious implications for large offshore wind farms, but are also a fascinating and sometimes problematic aspect of meteorological modelling. This thesis is an exploration of the predictability of mesoscale wind fluctuations from a statistical and physical modelling perspective. The motivation for the work was the large, hour-scale wind fluctuations that are sometimes observed at the Horns Rev I wind farm in the Danish North Sea. In particular, the thesis seeks to assess the practical possibility of predicting mesoscale wind fluctuations on a day-ahead time frame, since this is an important time scale for wind power spot market forecasts and decisions about reserve power allocation.

11.1 Summary of the work

Mesoscale atmospheric variability falls in between the study of turbulence, where short segments of time series are often considered as statistically stationary, and the study of average wind conditions, where hour-scale fluctuations are irrelevant. An important observation about long time series of wind speed is that they are non-stationary on many different time scales. Therefore, a new methodology for the climatological analysis of mesoscale wind fluctuations has been proposed, where non-stationarity is taken into account through the adaptive Hilbert Huang Transform, and conditional spectra are calculated by binning and averaging the time axis of Hilbert spectrum. In this case, the conditioning factors were meteorological observations such as wind direction, air-sea temperature difference or pressure tendency, but the methodology has a more general application to non-stationary time series analysis.

The methodology for developing conditional climatologies of wind variability was applied to time series observations from the first two large offshore in the world, Horns Rev I and Nysted. The results revealed that there are favourable times of year for large amplitude wind fluctuations at the two sites, the most striking result being that wind fluctuations between 1 and 5 hours are enhanced at Horns Rev between mid-September and March. The results

also showed a strong tendency for large wind fluctuations at Horns Rev to occur in flow with a sea fetch in contrast to flow coming from the land. At Nysted, there were also favourable direction bands for large amplitude wind fluctuations, but the patterns were much less clear, probably because of the more complex arrangement of land and sea around the site. Other key results were that variability tended to be higher in strong wind speeds at both sites, that variability was higher when the pressure is quickly falling or rising at Horns Rev, and when the pressure is quickly rising at Nysted, and that variability was higher when the air temperature was more than about 2 degrees lower than the sea water temperature.

The climatological analysis of local wind conditions at Horns Rev was extended by classifying 'severe variability events' and average meteorological variables in terms of the large scale weather patterns, satellite pictures and wind classes. The three classification methods were different to one another, but all presented a consistent message. The analysis of large scale weather patterns, which was undertaken using self-organising maps, showed that large amplitude wind variability tended to occur in cold-air outbreaks. The satellite analysis suggested that many of the observed severe variability cases in 2003 occurred in the presence of open cellular convection, a phenomenon that is often observed with cold air outbreaks. The wind-classes analysis showed that severe wind variability occurred when the large-scale flow was from the westerly sector and when the Froude number was negative. All these results are also consistent with the climatological analysis based on local conditions. Results from the climatological and classification analysis at Horns Rev were combined to test the predictability of mesoscale wind variability based on statistical information from the large scale flow. The models showed some skill, but could only achieve a 15% improvement in RMSE on a persistence forecast.

A different approach to modelling mesoscale wind variability using statistical methods is to down-scale large scale weather models using a mesoscale numerical weather prediction model in which mesoscale wind fluctuations can be resolved explicitly. Eight cases of severe wind variability, together with a control case of very low wind variability were chosen for case study analysis using the mesoscale model WRF. In all eight cases, open cellular convection was observed in satellite pictures over the North Sea, and in all cases cellular patterns of the right length scale for open cellular convection developed in the WRF simulations that were run with a horizontal grid spacing of 2 km. In contrast, very little mesoscale wind variability developed in the simulated control case. Time series of wind speed from the model showed that realistic hour-scale wind fluctuations developed in most cases, even the cases where the mean wind was inaccurate. Mesoscale model output is often provided to end users in hourly time steps for practical reasons. A new metric was included in the WRF model that calculates the variance during each hour using a single-pass algorithm. The algorithm proved to be a useful way of spatially presenting areas of high wind variability without storing high frequency model data, and paves the way for the calculation of other on-line sub-hourly statistics such as integrated energy production or maximum hourly wind speed.

There are many factors that affect a mesoscale model simulation, including the large scale weather forcing, topographic effects and surface inhomogeneities such as sea surface temperature gradients. These factors confuse the analysis of a single mesoscale phenomena such as open cellular convection, since they cause spatial and temporal non-stationarity and can result in spectral blending of the phenomena of interest with other features of similar length or time scales. These problems were addressed as far as possible by running an idealised version of the WRF model that simulated the simple case of cold, unstable air blowing over comparatively warmer sea water. The model was an excellent framework both for testing

the settings in the WRF model itself, and for assessing the sensitivity of the modelled open cellular convection to small changes in atmospheric conditions. Although only a small subset of the parameter space in which open cellular convection can occur was tested, the results revealed that the development of open cellular convection is, in fact, very sensitive to the boundary layer height, the existence of a capping inversion and the sea surface temperature. These results show that the correct development of open cellular convection is dependent on the accurate initialisation of the model and sufficiently detailed sea surface temperature data.

The sensitivity of the modelled open cellular convection to small changes in the boundary layer height and the sea surface temperature raised interesting questions about the forcing and energetics of the phenomenon. These questions were explored in depth by calculating a cell-scale kinetic energy budget of the cells in the idealised simulation. The calculations followed several cases in the literature, where incomplete budgets for convective rolls or cells were presented based on observational evidence. The mesoscale simulations offered the advantage of a full, gridded data set which allowed, for the first time, calculation of every term in the budget. An important difference between a mesoscale model and the real atmosphere is the way that energy is removed from the system, and this means that the diffusive processes in the model had to be carefully taken into account to present a balanced budget. The major result from the calculation of cell-scale kinetic energy budget was that the buoyancy term was nearly balanced by the pressure-balance term, and that the third most important term was the inter-scale transfer. There is, of course, no real evidence that the energy budget of the simulated open cells is a good model for the energy budget of open cells in the real atmosphere. However, the analysis revealed interesting results about the way that energy is removed from the boundary layer, and laid the ground work for a follow-up study using detailed observations or LES modelling.

11.2 Implications and future extensions to the work

Although the development and implementation of an operational forecasting system for wind variability was not possible within the scope of this thesis, the building blocks for such a model are described: The statistical analysis of important explanatory variables, the testing of simple statistical models, the verification of mesoscale wind fluctuations in a mesoscale model and the implementation of derived model parameters that describe the sub-hour scale model variability. A practical, operational forecast of wind variability could be provided in several formats, depending on the needs of end users. For example, a forecast of frequency and amplitude of expected wind fluctuations could alert transmission system operators and wind farm owners to the type of reserve power needed to respond on the expected time scales. Or, a simpler but potentially more useful forecast could take the format of a warning system for severe wind fluctuations. In either sense, the forecasts of wind variability can be provided as supplementary information to the forecasts of a mean wind speed that does not contain mesoscale fluctuations itself.

Variability forecasting is linked to probabilistic forecasting, but is not the same. It is obvious that a forecast may have a low certainty, but that none of the likely forecast outcomes involve risky conditions for high wind variability. Conversely, there may be high uncertainty in a forecast regarding, for example, the exact position of a low pressure centre and the timing of a front, but it might be rather certain that all possible scenarios will lead to large wind

fluctuations. The existence of severe wind mesoscale variability undermines the hour-scale predictability of the wind speed. The application of probabilistic forecasting to the problems of predicting mesoscale wind variability is also an interesting area for future research.

Given the lack of phase information in day-ahead forecasts of mesoscale wind variability, purely deterministic time series forecasts of wind speed or power with a good representation of mesoscale wind fluctuations will not be useful at all. For example, if the fluctuations are out of phase with the observed fluctuations, then the forecast will have large errors at all times, despite the fact that the statistical properties of the forecast might be correct. This means that the proper use of forecasts of mesoscale wind fluctuations will benefit from a shift away from purely deterministic forecasts and towards forecasts that describe the stochastic or higher order properties of the wind speed. For example, forecasts of frequency and amplitude, or an overall variability risk index, could alert wind farm operators or transmission system operators to the right mix of reserve power requirements. Using higher order information about the wind speed does not just mean issuing day-ahead warnings about possible large wind fluctuations, but probably requires rethinking the way that wind power is considered in the electricity system. In a future power system with very high wind power penetration, one could imagine an automated system that could optimally predict and assign reserve electricity requirements, and even incorporate demand-side management with the necessary response times to balance the predicted wind power fluctuations.

Although the motivation for the project was the large fluctuations in generated power at large offshore wind farms, the thesis concentrates almost exclusively on variability in the wind speed itself. There are several reasons for this. Firstly, the non-linear nature of the power curve seriously contaminates any relationships between variability and atmospheric conditions, since the size of the power fluctuations depends both on the size of the wind speed fluctuations and the gradient of the power curve. The blending of the physical properties of the atmosphere with the technical properties of the wind turbines would make the types of analysis presented here almost impossible if they addressed fluctuations in wind power. However, for the creation of useful forecasts, the transformation between wind and power fluctuations is essential. The first step in such a transformation involves mapping wind speed fluctuations to power fluctuations using the power curve of the wind turbine or wind farm, but there are also important considerations relating to the smoothing of the fluctuations by groups of turbines, whole wind farms or groups of wind farms within a region. In other words, the spatial coherence of the wind speed time series in the mesoscale range of frequencies must be considered. In the context of organised mesoscale convection, the potential for highly correlated observations over a large spatial area is an interesting consideration in itself.

This thesis was about the predictability of mesoscale wind fluctuations from the perspective of short term (day-ahead) forecasting. However, there is another important aspect to the predictability of mesoscale wind fluctuations that was not addressed here. This is the climatological impact of wind fluctuations on the long term financial profitability of large offshore wind farms. That is, how much does the frequent occurrence of large hour-scale wind fluctuations degrade the value of wind farms, and in particular large offshore wind farms? This is a complicated question, the answers to which span the economic analysis of the market integration of large scale wind power, the technical issues relating to integration of wind power into the grid, the predictability of large wind fluctuations using meteorological and statistical modelling, the spatial coherence of mesoscale wind fluctuations over large areas with high wind penetration and the climatological analysis of the occurrence of large wind fluctuations. Although there are too many unknowns in all of these areas to provide definite

answers about the financial cost of mesoscale wind fluctuations, it is essential that these questions are at least considered when we discuss the problems of wind variability. Many of the results presented in this thesis are also highly relevant to the assessment of potential wind farm sites in terms of the mesoscale wind variability.

In terms of mesoscale modelling, this project focused on open cellular convection, which was shown to be associated with many of the severe wind variability events at Horns Rev I. Based on the analysis of short 1 and 4 year time series in this thesis, open cellular convection was shown to be the dominant phenomenon associated with severe wind variability. However, there are clearly other meteorological process that lead to mesoscale wind variability, including gravity waves and convective rolls, as well as scattered convection, thunderstorms, sea-breeze circulations and low pressure troughs. It was not possible to thoroughly analyse all of these phenomena, but the detailed study of open cellular convection demonstrates, in principle, the use of mesoscale models for forecasting mesoscale wind variability. The analysis of open cellular convection in WRF was based closely on one set of atmospheric conditions, from which small deviations in boundary layer height and sea surface temperature were made. A broader study, probably based on climatological analysis of real observed cases of open cellular convection and long model runs, could further explore the relationship between cell size (and therefore frequency of wind fluctuations), cell intensity (and therefore the amplitude of wind fluctuations) and the atmospheric conditions.

The predictability of mesoscale wind fluctuations is not only an important consideration for the end users of mesoscale modelling results, but is important for mesoscale model development itself. For example, mesoscale shallow convection transfers energy from the surface to the upper atmosphere, and so it is important that it is modelled accurately, or that its effects are properly parametrised in the model. The results presented in this thesis showed that open cellular convection is predictable using a mesoscale model and that the modelled open cellular convection is highly sensitive to small changes in the atmospheric conditions. However, the extent to which the energy transfer, the entrainment at the top of the boundary layer and the plumes overshooting the top of the boundary layer are realistic is unknown, and since mesoscale shallow convection covers a large portion of the world's oceans, this is an important area for future model development. The calculation of the cell-scale kinetic energy budget provided a fascinating insight into these questions and into the sub-grid-scale diffusion in the model. However, the budget calculated here was only based on a limited number of cases, and the unanswered questions regarding the impact of horizontal grid resolution and the choice of boundary layer parametrisation on the resolved and sub-grid-scale fluctuations pose interesting areas for future research.

As computing power increases, the possibility of using very high resolution models for operational purposes will become more and more realistic. However, running very high resolution models is not simply a matter of decreasing the horizontal grid spacing, because the planetary boundary layer scheme becomes redundant at some unknown resolution and explicit vertical mixing is possible. Producing weather forecasts using LES models is an exciting and tempting area, but there are many important areas of research to be addressed first. For example, reproducing the idealised simulations of open cellular convection presented here using LES modelling would be an important first step to assessing the mesoscale wind fluctuations in an LES model in comparison to those in mesoscale models and those observed in the real world.

In the future, the ever increasing amount of wind power generation will continue to present

interesting and important challenges relating to the predictability and variability of the wind speed. It is expected that these challenges will be met through a combination of statistical and physical modelling, and through an understanding of the meteorological processes that lead to certain patterns in wind speed.

Bibliography

Ackermann, T., J. R. Abbad, I. M. Dudurych, I. Erlich, H. Holttinen, J. R. Kristoffersen, and P. E. Sørensen, 2007: Technical aspects of status and expected future trends for wind power in Denmark. *Wind Energy*, **10**, 31–49.

Agee, E. M., 1987: Meso-scale cellular convection over the oceans. *Dyn. Atmos. Oceans*, **10**, 317–341.

Agee, E. M. and T. S. Chen, 1973: A model for investigating eddy viscosity effects on mesoscale cellular convection. *J. Atmos. Sci.*, **30**, 180–189.

Agee, E. M., T. S. Chen, and K. E. Dowell, 1973: A review of mesoscale cellular convection. *Bull. Amer. Meteor. Soc.*, **54**, 1004–1012.

Agee, E. M. and K. E. Dowell, 1974: Observational studies of mesoscale cellular convection. *J. Appl. Meteor.*, **13**, 46–53.

Akhmatov, V., C. Rasmussen, P. B. Eriksen, and J. Pedersen, 2007: Technical aspects of status and expected future trends for wind power in Denmark. *Wind Energy*, **10**, 31–49.

Atkinson, B. W. and J. W. Zhang, 1996: Mesoscale shallow convection in the atmosphere. *Rev. Geophys.*, **34**, 403–431.

Badger, J., G. Giebel, X. G. Larsen, T. S. Nielsen, H. A. Nielsen, H. Madsen, and J. Tøfting, 2007: Report on the use of stability parameters and mesoscale modelling in short-term prediction. Technical Report Risø-R-1614, Risø National Laboratory, Technical University of Denmark, Roskilde, Denmark.

Bakan, S., 1985: The structure of open cellular convection as revealed by time series of surface obesrvations: a case study. *Contrib. Atmos. Phys.*, **58**, 11–16.

Bakan, S. and E. Schwarz, 1992: Cellular convection over the North-Eastern Atlantic. *Int. J. Climatol.*, **12**, 353–367.

Barthelmie, R. J., K. Hansen, S. T. Frandsen, O. Rathmann, J. G. Schepers, W. Schlez, J. Phillips, K. Rados, A. Zervos, E. S. Politis, and P. K. Chaviaropoulos, 2009: Modelling and measuring flow and wind turbine wakes in large wind farms offshore. *Wind Energy*, **12**, 431–444, doi:10.1002/we.348.

Barthlott, C., P. Drobinski, C. Fesquet, T. Dubos, and C. Pietras, 2007: Long-term study of coherent structures in the atmospheric surface layer. *Bound. Layer Meteorol.*, **125**, 1–24.

Bennett, L. J., T. M. Weckwerth, A. M. Blyth, B. Geerts, Q. Miao, and Y. P. Richardson, 2010: Observations of the evolution of the nocturnal and convective boundary layers and the structure of open-celled convection on 14 June 2002. *Mon. Wea. Rev.*, **138**, 2589–2607, doi:10.1175/2010MWR3200.1.

Bierbooms, W., 2005: Investigation of spatial gusts with extreme rise time on the extreme loads of pitch-regulated wind turbines. *Wind Energy*, **8**, 17–34.

Boashash, B., 1992: Estimating and interpreting the instantaneous frequency of a signal, Part 1: Fundamentals. *Proc. IEEE*, **80**, 520–538.

Box, G. E. P. and G. M. Jenkins, 1970: *Time series analysis forecasting and control*. Holden-Day, 553 pp.

Brown, R. A., 1980: Longitudinal instabilities and secondary flows in the planetary boundary layer: A review. *Rev. Geophys. Space Phys.*, **18**, 683–697.

Brümmer, B., 1985: Structure, dynamics and energetics of boundary layer rolls from KonTur aircraft observations. *Contrib. Atmos. Phys.*, **58**, 237–254.

— 1999: Roll and cell convection in wintertime arctic cold-air outbreaks. *J. Atmos. Sci.*, **56**, 2613–2636.

Brümmer, B., T. Fisher, and S. Zank, 1986: Aircraft observations of open cellular structures during KonTur. *Contrib. Atmos. Phys.*, **59**, 162–184.

Brusch, S., S. Lehner, and J. Schulz-Stellenfleth, 2008: Synergetic use of radar and optical satellite images to support severe storm prediction for offshore wind farming. *IEEE J. Earth Obs. Remot. Sens.*, **1**, 57–66.

Busack, B., S. Bakan, and H. Luthardt, 1985: Surface conditions during mesoscale cellular convection. *Contrib. Atmos. Phys.*, **58**, 4–10.

Casado, M. J., M. A. Pastor, and F. J. Doblas-Reyes, 2009: Euro-Atlantic circulation types and modes of variability in Winter. *Theor. Appl. Climatol.*, **96**, 17–29.

Chambers, J. M. and T. J. Hastie, 1992: *Statistical models in S*. Wadsworth & Brooks / Cole advanced books and software, 608 pp.

Christiansen, M. B. and C. B. Hasager, 2005: Wake effects of large offshore wind farms identified from satellite SAR. *Remote. Sens. Environ.*, **98**, 251–268.

— 2006: Using airborne and satellite SAR for wake mapping offshore. *Wind Energy*, **9**, 437–455.

Christiansen, M. B., W. Koch, J. Horstmann, C. B. Hasager, and M. Nielsen, 2006: Wind resource assessment from C-band SAR. *Remote. Sens. Environ.*, **105**, 68–81.

Cieszelski, R., 1998: A case study of Rayleigh-Bénard convection with clouds. *Bound. Layer Meteorol.*, **88**, 211–237.

Coelingh, J., A. van Wijk, and A. Holtslag, 1996: Analysis of wind speed observations over the North Sea. *J. Wind Eng. Ind. Aerodyn.*, **61**, 51–69.

— 1998: Analysis of wind speed observations over the North Sea coast. *J. Wind Eng. Ind. Aerodyn.*, **73**, 125–144.

Coelingh, J., A. J. M. van Wijk, J. W. Cleijne, and R. Pleune, 1992: Description of the North Sea wind climate for wind energy applications. *J. Wind Eng. Ind. Aerodyn.*, **39**, 221–232.

Cohen, L., 1989: Time-frequency distributions – a review. *Proc. IEEE*, **77**, 941–981.

— 1995: *Time Frequency Analysis*. Prentice-Hall, 320 pp.

Costa, A., A. Crespo, J. Navarro, G. Lizcano, H. Madsen, and E. Feitosa, 2008: A review on the young history of the wind power short-term prediction. *Ren. Sust. Energy Rev.*, **12**, 1725–1744.

Cuell, C. and B. Bonsal, 2009: An assessment of climatological synoptic typing by principal component analysis and k-means clustering. *Theor. Appl. Climatol.*, **98**, 361–373.

Davey, R., M. Woods, C. Russell, and P. Coppin, 2010: Statistical downscaling of wind variability from meteorological fields. *Bound. Layer Meteorol.*, **135**, 161–175.

de Bodt, E., M. Cottrell, and M. Verleysen, 2002: Statistical tools to assess the reliability of self-organising maps. *Neural Networks*, **15**, 967–978.

de Roode, S. R., P. G. Duynkerke, and H. J. J. Jonker, 2004: Large-eddy simulation - how large is large enough? *J. Atmos. Sci.*, **61**, 403–420.

Draxl, C., A. N. Hahmann, A. Pe na, J. N. Nissen, and G. Giebel, 2010: Validation of boundary layer winds from WRF mesoscale forecasts with applications to wind energy forecasting. *Proceedings*, Keystone, CO, USA.

Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107.

Duffy, D., 2004: The application of Hilbert-Huang transforms to meteorological datasets. *J. Atmos. Oceanic Technol.*, **21**, 599–611.

Errico, R. M., 1985: Spectra computed from a limited area grid. *Mon. Wea. Rev.*, **113**, 1554–1562.

Fealy, R. and J. Sweeney, 2007: Identification of frequency changes in synoptic circulation types and consequences for glacier mass balance in Norway. *Norwegian J. Geog.*, **61**, 76–91.

Focken, U., M. Lange, M. Monnich, H. Waldl, H. Beyer, and A. Luig, 2002: Short-term prediction of the aggregated power output of wind farms - a statistical analysis of the reduction of the prediction error by spatial smoothing effects. *J. Wind Eng. Ind. Aerodyn.*, **90**, 231–246.

Gabor, D., 1946: Theory of communication. *J. IEE*, **93**, 429–457.

Giebel, G., 2000: On the benefits of distributed generation of wind energy in Europe. PhD thesis, Carl von Ossietzky Universit:at, Oldenberg, Germany.

— 2011: The state-of-the-art in short-term prediction of wind power - a literature overview, 2nd edition. Deliverable report d-1.2, ANEMOS.plus.

Giraudel, J. L. and S. Lek, 2001: A comprison of self-organizing map algorithm and some conventional methods for ecological community coordination. *Ecological Modelling*, **146**, 329–339.

Gjerstad, J., S. E. Aasen, H. I. Andersson, I. Brevik, and J. Løvseth, 1995: An analysis of low-frequency maritime atmospheric turbulence. *J. Atmos. Sci.*, **52**, 2663–2669.

Grachev, A. A. and C. W. Fairall, 1997: Dependence of the Monin-Obukhov stability parameter on the bulk Richardson number of the ocean. *J. Appl. Meteor.*, **36**, 406–414.

Hahn, R. S. and C. F. Mass, 2009: The impact of positive definite moisture advection and low-level moisture flux bias over orography. *Mon. Wea. Rev.*, **137**, 3055–3071.

Heinle, A., A. Macke, and A. Srivastav, 2010: Automatic cloud classification of whole sky images. *Atmos. Meas. Tech.*, **3**, 557–567.

Helfand, M. and E. Kalnay, 1983: A model to determine open or closed cellular convection. *J. Atmos. Sci.*, **40**, 631–650.

Hering, A. S. and M. G. Genton, 2010: Powering up with space-time wind forecasting. *J. Am. Stat. Assoc.*, **105**, 92–104, doi:10.1198/jasa.2009.ap18117.

Hewitson, B. C. and R. G. Crane, 2002: Self-organizing maps: applications to synoptic climatology. *Clim. Res.*, **22**, 13–26.

Hinzpeter, H., 1985: KonTur-results (foreword). *Contrib. Atmos. Phys.*, **58**, 1–3.

Holttinen, H., 2005: Optimal electricity market for wind power. *Energy Policy*.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341.

Huang, N., 2005a: *Computing instantaneous frequency by normalizing Hilbert transform*. U. S. Patent Office, patent number US 6,901,353.

— 2005b: *Hilbert-Huang Transform and its applications*. World Scientific Publishing Company, 1-26 pp.

Huang, N., M. C. Wu, S. R. Long, S. S. P. Shen, W. Qu, P. Gloersen, and K. L. Fan, 2003: A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. *Proc. Roy. Soc. A.*, **459**, 2317–2345.

Huang, N. and Z. Wu, 2008: A review on Hilbert-Huang transform: Method and its applications to geophysical studies. *Rev. Geophys.*, **46**, RG2006.

Huang, N., Z. Wu, S. Long, K. Arnold, X. Chen, and K. Blank, 2009: On instantaneous frequency. *Adv. Adaptive Data Anal.*, **1**, 177–229.

Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, 1998: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. Roy. Soc. A.*, **454**, 903–995.

Hundsdorfer, W., B. Koren, M. van Loon, and J. G. Verwer, 1995: A positive finite-difference advection scheme. *J. Comput. Phys.*, **117**, 35–46.

James, P. M., 2007: An objective classification method for Hess and BrezowskyGrosswetterlagen over Europe. *Theor. Appl. Climatol.*, **88**, 17–42.

Kain, J. S. and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2784–2802.

Kim, D., S.-H. Paek, and H.-S. Oh, 2008: A Hilbert-Huang transform approach for prediction cyber-attacks. *J. Korean Stat. Soc.*, **37**, 277–283.

Knievel, J. C., G. H. Bryan, and J. P. Hacker, 2007: Explicit numerical diffusion in the WRF model. *Mon. Wea. Rev.*, **135**, 3808–3824.

Kohonen, T., 2001: *Self-organizing maps*. Springer series in information sciences, Springer-Verlag, Berlin, 3rd edition.

Krishnamurti, R., 1975: On cellular cloud patterns. Part 1: Mathematical model. *J. Atmos. Sci.*, **32**, 1353–1363.

Krueger, A. F. and S. Fritz, 1961: Cellular cloud patterns revealed by Tiros I. *Tellus*, **13**, 1–7.

Kusiak, A., W. Li, and Z. Song, 2010: Dynamic control of wind turbines. *Ren. Energy*, **35**, 456–463.

Labat, D., 2005: Recent advances in wavelet analysis: Part 1. a review of concepts. *J. Hydrol.*, **314**, 275–288.

Landberg, L., 1997: The availability and variability of the European wind resource. *Int. J. Solar Energy*, **18**, 313–320.

— 1999: Short-term prediction of the power production from wind farms. *J. Wind Eng. Ind. Aerodyn.*, **80**, 207–220.

— 2001: Short-term prediction of local wind conditions. *J. Wind Eng. Ind. Aerodyn.*, **89**, 235–245.

Lange, B., H. Waldl, A. G. Guerrero, D. Heinemann, and R. J. Barthelmie, 2003: Modelling of offshore wind turbine wakes with the wind farm program FLaP. *Wind Energy*, **6**, 68–104.

Lange, M., U. Focken, and A. Lenz, 2010: Ramp event forecasting - practical experiences in the USA and Australia. *4th workshop on best practice in the use of short term forecasting of wind power*, Québec City, Canada.

Larsen, G. C., J. Mann, T. A. Ighil, and A. S. Mouritzen, 2009: Do wind farms influence large scale turbulence? Extended Abstracts for Euromech Colloquium 508, 38–40.

Larsén, X. G. and S. E. Larsen, 2010: A case study of mesoscale spectra of wind and temperature, observed and simulated. *Quart. J. Roy. Meteor. Soc.*, in press.

LeMone, M. A., 1973: The structure and dynamics of horizontal roll vortices in the planetary boundary layer. *J. Atmos. Sci.*, **30**, 1077–1091.

— 1976: Modulation of turbulence energy by longitudinal rolls in an unstable planetary boundary layer. *J. Atmos. Sci.*, **33**, 1308–1320.

Liu, Y., R. H. Weisberg, and C. N. K. Mooers, 2006: Performance evaluation of the self-organising map for feature extraction. *J. Geophys. Res.*, **11**, doi:10.1029/2005JC003117.

Louis, J. F., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Bound. Layer Meteorol.*, **17**, 187–202.

Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.*, **83**, 407–430.

McCullagh, P. and J. A. Nelder, 1989: *Generalised Linear Models*. Cambridge University Press, 2nd edition, 608 pp.

Mellor, G. L. and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space. Phys.*, **20**, 851–875.

Michaelides, S. C., F. Liassidou, and C. N. Schizas, 2007: Synoptic classification and establishment of analogues with artificial neural networks. *Pure Appl. Geophys.*, **164**, 1347–1364.

Miura, Y., 1986: Aspect ratios of longitudinal rolls and convection cells observed during cold air outbreaks. *J. Atmos. Sci.*, **43**, 26–39.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res. Atmos.*, **102**, 16663–16682.

Moeng, C.-H., J. Dudhia, J. Klemp, and P. Sullivan, 2007: Examining two-way grid nesting for large eddy simulation of the PBL using the WRF model. *Mon. Wea. Rev.*, **135**, 2295–2311.

Möhrlen, C., 2004: Uncertainty in wind energy forecasting. PhD Thesis, University College Cork, National University of Ireland, Cork, Ireland.

Molinari, J. and M. Dudek, 1992: Parameterization of convective precipitation in mesoscale numerical models: a critical review. *Mon. Wea. Rev.*, **120**, 326–344.

Moyer, K. A. and G. S. Young, 1994: Observations of mesoscale cellular convection from the marine stratocumulus phase of FIRE. *Bound. Layer Meteorol.*, **71**, 109–133.

Müller, G. and A. Chlond, 1996: Three-dimensional numerical study of cell broadening during cold-air outbreaks. *Bound. Layer Meteorol.*, **81**, 289–323.

Nakanishi, M. and H. Niino, 2009: Development of an improved turbulence closure model for the atmospheric boundary layer. *JMSJ*, **87**, 895–912.

Nielsen, T. S., 1999: Using meteorological forecast in on-line prediction of wind power. Technical report, Institute of mathematical modelling - Technical University of Denmark, Eltra, Elsam, Lyngby, Denmark.

Orlanski, I., 1975: A rational subdivision of scales for atmospheric motion. *Bull. Amer. Meteor. Soc.*, **56**, 527–530.

Painemal, D., R. Garreaud, J. Rutllant, and P. Zuidema, 2010: Southeast Pacific stratocumulus: High-frequency variability and mesoscale structures over San Félix Island. *J. Appl. Meteor. Climatol.*, **49**, 463–477.

Pankiewicz, G. S., 1997: New METEOSAT pattern recognition products for use in weather forecasting. *Adv. Space Res.*, **19**, 443–446.

Park, Y.-S., J. Tison, S. Lek, J.-L. Giraudel, M. Coste, and F. Delmas, 2006: Application of a self-organising map to select representative species in multivariate analysis: A case study determining diatom distribution patterns across France. *Ecological Informatics*, **1**, 247–257.

Peña, A., S.-E. Gryning, and C. B. Hasager, 2008: Measurements nad modelling of the wind speed profile in the marine atmospheric boundary layer. *Bound. Layer Meteorol.*, **129**, 479–495.

Philipp, A., J. Bartholy, C. Beck, M. Erpicum, P. Esteban, X. Fettweis, R. Huth, P. James, S. Jourdain, F. Kreienkamp, T. Krennert, S. Lykoudis, S. C. Michalides, K. Pianko-Kluczynska, P. Post, D. Rasilla Alvarez, R. Schiemann, A. Spekat, and F. S. Tymvios, 2010: Cost733cat-A database of weather and circulation type classifications. *Phys. Chem. Earth*, **35**, 360–373.

Pielke sr., R. A., 2002: *Mesoscale meteorological modelling*, volume 78 of *International Geophysics Series*. Academic Press, London, England, 2rd edition.

Pinson, P., C. Chevallier, and G. N. Kariniotakis, 2007: Trading wind generation from short-term probabilistic forecasts of wind power. *Power Systems, IEEE Transactions on*, **22**, 1148 –1156, doi:10.1109/TPWRS.2007.901117.

Pinson, P. and H. Madsen, 2010: Adaptive modelling and forecasting of wind power fluctuations with Markov-switching autoregressive models. *J. Forecasting*, doi:10.1002/for.1194.

Potter, C. W., E. Grimit, and B. Nijssen, 2009: Potential benefits of a dedicated probabilistic rapic ramp event forecast tool. *IEEE/PES power systems conference and exposition*, volume 1-3, 409–413.

Prudhomme, C. and M. Genevier, 2010: Can atmospheric circulation be linked to flooding in Europe? *Health Phys.*, doi:10.1002/hyp.7879.

Pryor, S. C. and R. J. Barthelmie, 2003: Long-term trends in near-surface flow over the Baltic. *Int. J. Climatol.*, **23**, 271–289.

— 2010: Climate change impacts on wind energy: A review. *Ren. Sust. Energy Rev.*, **14**, 430–437.

Ray, D., 1986: Variable eddy diffusivities and atmospheric cellular convection. *Bound. Layer Meteorol.*, **36**, 117–131.

Richardson, M. I., A. D. Toigo, and C. E. Newman, 2007: PlanetWRF: A general purpose, local to global numerical model for planetary atmospheric and climate dynamics. *J. Geophys. Res.*, **112**, doi:10.1029/2006JE002825.

Romero, R., G. Sumner, C. Ramix, and A. Genovés, 1999: A classification of the atmospheric circulation patterns producing significant daily rainfall in the Spanish Mediterranean area. *Int. J. Climatol.*, **19**, 765–785.

Sammon, jr., J. W., 1969: A nonlinear mapping for data structure analysis. *IEEE Trans. Computers.*, **C-18**, 401–409.

Schröter, M., S. Raasch, and H. Jansen, 2005: Cell broadening revisited: Results from high-resolution large-eddy simulation of cold air outbreaks. *J. Atmos. Sci.*, **62**, 2023–2032.

Shahin, M. A., H. R. Maier, and M. B. Jaksa, 2004: Data division for developing neural networks applied to geotechnical engineering. *J. Computing in Civil Eng.*.

Sikora, T. D., G. S. Young, C. M. Fisher, and M. Stepp, 2010: A synthetic aperture radar-based climatology of open cell convection over the northeast Pacific Ocean. *J. Appl. Meteor. Climatol.*, doi:10.1175/2010JAMC2624.1.

Simmons, A., S. Uppala, D. Dee, and S. Kobayashi, 2006/2007: Era-interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF Newsletter*, **110**, 25–35.
URL http://www.ecmwf.int/publications/newsletters/pdf/110_rev.pdf

Skamarock, W. C., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Wea. Rev.*, **132**, 3019–3032.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X.-Y. Huang, W. Wang, and J. G. Powers, 2008: A description of the advanced research WRF version 3. NCAR Technical Note TN-475+STR, NCAR, Boulder, Colorado.

Socal, G., F. Acri, M. Bastianini, F. B. Aubry, F. Bianchi, D. Cassin, J. Coppola, A. D. Lazzari, V. Bendelj, G. Cossarini, and C. Solidoro, 2008: Hydrological and biogeochemical features of the Northern Adriatic Sea in the period 2003–2006. *Marine Ecology*, **29**, 449–468.

Sørensen, J. D., S. Frandsen, and N. J. Tarp-Johansen, 2008a: Effective turbulence models and fatigue reliability in wind farms. *Prob. Eng. Mech.*, **23**, 531–538.

Sørensen, P., N. A. Cutululis, A. Vigueras-Rodríguez, H. Madsen, P. Pinson, L. Jensen, J. Hjerrild, and M. Donovan, 2008b: Modelling of power fluctuations from large offshore wind farms. *Wind Energy*, **11**, 29–43.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology*. Kluwer Academic Publishers, 666 pp.

Sweeney-Reed, C. M. and S. J. Nasuto, 2007: A novel approach to the detection of synchronisation in EEG based on empirical mode decomposition. *J. Comput. Neurosci.*, **23**, 79–111.

Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.*, **132**, 519–542.

Tuohy, A., P. Meibom, E. Denny, and M. O'Malley, 2009: Unit commitment for systems with significant wind penetration. *IEEE Trans. Power Syst.*, **24**, 592–601.

Venables, W. N. and B. D. Ripley, 2002: *Modern applied statistics with S*. Springer-Verlag, 495 pp.

Vesanto, J., 2000: *Using SOM in Data Mining*. Ph.D. thesis, Helsinki university of Technology, Espoo, Finland.

Vincent, C., C. Draxl, G. Giebel, P. Pinson, J. Jørgensen, and C. Möhrlen, 2009: Spectral verification of a mesoscale ensemble. *Proceedings*, European Wind Energy Conference and Exhibition, Marseille.

Vincent, C., G. Giebel, P. Pinson, and H. Madsen, 2010a: Resolving non-stationary spectral information in wind speed time series using the Hilbert-Huang transform. *J. Appl. Meteor. Climatol.*, **49**, 253–267.

Vincent, C., P. Pinson, and G. Giebel, 2010b: Wind fluctuations over the North Sea. *Int. J. Climatol.*, DOI: 10.1002/joc.2175.

Vincent, C. L., M. C. Kelly, and A. N. Hahmann, 2011: Idealised WRF simulations of open cellular convection over the sea. Submitted to Boundary Layer Meteorology.

Wang, H. and G. Feingold, 2009a: Modelling cellular structures and drizzle in marine stratocumulus. Part I: Impact of drizzle on the formation and evolution of open cells. *J. Atmos. Sci.*, **66**, 3237–3256.

— 2009b: Modelling cellular structures and drizzle in marine stratocumulus. Part II: The microphysics and dynamics of the boundary region between open and closed cells. *J. Atmos. Sci.*, **66**, 3257–3275.

Wang, W., C. Bruyère, M. Duda, J. Dudhia, D. Gill, H.-C. Lin, J. Michalades, S. Rizvi, and X. Zhang, 2010: Weather research and forecasting ARW version 3 modelling system user's guide. User's guide, NCAR, Boulder, USA.

Wehrens, R. and L. Buydens, 2007: Self- and super-organising maps in R: the kohonen package. *J. Stat. Softw.*, **21**.
URL http://www.jstatsoft.org/v21/i05

Welford, B. P., 1961: Note on a method for calculating corrected sums of squares and products. *Technometrics*, **4**, 419–420.

Weusthoff, T. and T. Hauf, 2008: The life cycle of convective-shower cells under post frontal conditions. *Quart. J. Roy. Meteor. Soc.*, **134**, 841–857.

Wicker, L. J. and W. C. Skamarock, 2002: Time-splitting methods for elastic models using forward time series. *Mon. Wea. Rev.*, **130**, 2088–2097.

Wu, Z. and N. Huang, 2009: Ensemble empirical mode decomposition: A noise assisted data analysis method. *Adv. Adaptive Data Anal.*, **1**, 1–41.

Xingjie, L., M. Zengqiang, L. Peng, and M. Huawei, 2009: Study on the multi-step forecasting for wind speed based on EMD. *1st international conference on sustainable power generation and supply, SUPERGEN '09*, Nanjing, China.

Yu, L., S. Want, and K. K. Lai, 2008: Forecasting crude oil price with an emd-based neural network ensemble learning paradigm. *Energy Economics*, **30**, 2623–2635.

Zdunkowski, W. and A. Bott, 2003: *Dymamics of the atmosphere - a course in theoretical meteorology*. Cambridge University Press, Cambridge, United Kindom.

Appendix A

Derivation of algorithm for updating variance

Updating mean

$$\bar{x}_n = \bar{x}_{n-1} + \frac{1}{n}(x_n - \bar{x}_{n-1}) \tag{A.1}$$

Updating sum of squares

$$\begin{aligned}
P_n &= \sum_{i=1}^{n}(x_i - \bar{x}_n)^2 \\
&= \sum_{i=1}^{n}((x_i - \bar{x}_{n-1}) - \frac{1}{n}(x_n - \bar{x}_{n-1}))^2 \\
&= \sum_{i=1}^{n-1}(x_i - \bar{x}_{n-1})^2 - 2(x_n - \bar{x}_{n-1})\frac{1}{n}\sum_{i=1}^{n} x_i + 2\bar{x}_{n-1}x_n - 2\bar{x}_{n-1}^2 + \frac{1}{n}(x_n - \bar{x}_{n-1})^2 \\
&= P_{n-1} + (x_n - \bar{x}_{n-1})(x_n + \bar{x}_{n-1} + \frac{1}{n}(x_n - \bar{x}_{n-1}) - 2\bar{x}_n) \\
&= P_{n-1} + (x_n - \bar{x}_{n-1})(x_n - \bar{x}_n)
\end{aligned} \tag{A.2}$$

Updating Variance

$$\sigma_n^2 = \frac{P_n}{n-1} \tag{A.3}$$

Appendix B

Derivation of the cell-scale kinetic energy budget

The derivation begins with the Navier Stokes equation:

$$\frac{\partial \mathbf{V}}{\partial t} = -(\mathbf{V}\cdot\nabla)\mathbf{V} - \frac{1}{\rho}\nabla p - 2\Omega_e \times \mathbf{V} + g\frac{\theta_v}{\theta_{v,0}}\mathbf{k} \tag{B.1}$$

The cell scale kinetic energy equation is obtained by multiplying the Navier Stokes equation by the cell-scale velocity, $\mathbf{v_l}$:

$$\underbrace{\mathbf{v_l}\cdot\frac{\partial \mathbf{V}}{\partial t}}_{\text{TERM 1}} = -\underbrace{\mathbf{v_l}\cdot(\mathbf{V}\cdot\nabla)\mathbf{V}}_{\text{TERM 2}} - \underbrace{\frac{1}{\rho}\mathbf{v_l}\cdot\nabla p}_{\text{TERM 3}} - \underbrace{\mathbf{v_l}\cdot 2\Omega_e\times\mathbf{V}}_{\text{TERM 4}} + \underbrace{\mathbf{v_l}\cdot g\frac{\theta_v}{\theta_{v,0}}\mathbf{k}}_{\text{TERM 5}} \tag{B.2}$$

The cell scale kinetic energy equation is filtered by averaging over the domain. The $\widehat{\ }$ denotes the domain average.

$$\underbrace{\widehat{\mathbf{v_l}\cdot\frac{\partial \mathbf{V}}{\partial t}}}_{\text{TERM 1}} = -\underbrace{\mathbf{v_l}\cdot\widehat{(\mathbf{V}\cdot\nabla)\mathbf{V}}}_{\text{TERM 2}} - \underbrace{\widehat{\frac{1}{\rho}\mathbf{v_l}\cdot\nabla p}}_{\text{TERM 3}} - \underbrace{\mathbf{v_l}\cdot\widehat{2\Omega_e\times\mathbf{V}}}_{\text{TERM 4}} + \underbrace{\mathbf{v_l}\cdot\widehat{g\frac{\theta_v}{\theta_{v,0}}\mathbf{k}}}_{\text{TERM 5}} \tag{B.3}$$

For convenience, hats will be omitted in the following derivation, but it can be assumed that there is a hat over every individual term. The velocity, pressure and virtual potential temperature are decomposed into a mean term, A_∞, a cell-scale term, a_l and a turbulent fluctuations term, a':

$$A = A_\infty + a_l + a'$$

The decomposition is applied to all the terms in equation B.3, together with the averaging rules,

$$\widehat{a'} = 0, \widehat{a_l} = 0, \widehat{a'b_l} = 0, \widehat{a'b_\infty} = 0, \frac{\partial}{\partial x_i}\widehat{a'} = 0, \frac{\partial}{\partial x_i}\widehat{a'b_l} = 0, \frac{\partial}{\partial x_i}\widehat{a'b_\infty} = 0,$$

where a and b are any of the three velocity components, pressure or virtual potential temperature, and the assumptions,

$$\frac{\partial U_\infty}{\partial x} = \frac{\partial U_\infty}{\partial y} = \frac{\partial V_\infty}{\partial x} = \frac{\partial V_\infty}{\partial y} = 0, \quad W_\infty = 0, \quad \frac{\partial U_\infty}{\partial t} = \frac{\partial V_\infty}{\partial t} = \frac{\partial W_\infty}{\partial t} = 0.$$

Term 1

$$\begin{aligned}
\mathrm{T1} &= \mathbf{v_l} \cdot \frac{\partial \mathbf{V}}{\partial t} \\
&= \frac{\partial}{\partial t}(\frac{1}{2}\mathbf{v_l^2}) + \mathbf{v_l} \cdot \frac{\partial \mathbf{V}_\infty}{\partial t} + \mathbf{v_l} \cdot \frac{\partial \mathbf{v}'}{\partial t} \\
&= \frac{\partial}{\partial t}(\frac{1}{2}\mathbf{v_l^2}), \text{ since } \frac{\partial U_\infty}{\partial t} = \frac{\partial V_\infty}{\partial t} = \frac{\partial W_\infty}{\partial t} = 0 \text{ and } \widehat{a_l \frac{\partial b'}{\partial t}} = 0. \\
&= \frac{\partial E_l}{\partial t} \\
&\text{where } E_l = \frac{1}{2}(u_l^2 + v_l^2 + w_l^2)
\end{aligned} \tag{B.4}$$

Term 2

$$\begin{aligned}
\mathrm{T2} &= \mathbf{v_l} \cdot (\mathbf{V} \cdot \nabla)\mathbf{V} \\
&= \mathbf{v_l} \cdot ((\mathbf{V}_\infty + \mathbf{v_l} + \mathbf{v}') \cdot \nabla)(\mathbf{V}_\infty + \mathbf{v_l} + \mathbf{v}') \\
&= \mathbf{v_l} \cdot ((\mathbf{V}_\infty \cdot \nabla)\mathbf{V}_\infty + (\mathbf{V}_\infty \cdot \nabla)\mathbf{v_l} + (\mathbf{V}_\infty \cdot \nabla)\mathbf{v}' \\
&\quad + (\mathbf{v_l} \cdot \nabla)\mathbf{V}_\infty + (\mathbf{v_l} \cdot \nabla)\mathbf{v_l} + (\mathbf{v_l} \cdot \nabla)\mathbf{v}' \\
&\quad + (\mathbf{v}' \cdot \nabla)\mathbf{V}_\infty + (\mathbf{v}' \cdot \nabla)\mathbf{v_l} + (\mathbf{v}' \cdot \nabla)\mathbf{v}')
\end{aligned} \tag{B.5}$$

The nine sub-terms in Term 2 are treated in the order that they appear in equation B.5, and are labelled T2.1 to T2.9 respectively.

$$\begin{aligned}
\mathrm{T2.1} &= \mathbf{v_l} \cdot (\mathbf{V}_\infty \cdot \nabla)\mathbf{V}_\infty \\
&= \mathbf{v_l} \cdot \begin{bmatrix} U_\infty \frac{\partial U_\infty}{\partial x} + V_\infty \frac{\partial U_\infty}{\partial y} + W_\infty \frac{\partial U_\infty}{\partial z} \\ U_\infty \frac{\partial V_\infty}{\partial x} + V_\infty \frac{\partial V_\infty}{\partial y} + W_\infty \frac{\partial V_\infty}{\partial z} \\ U_\infty \frac{\partial W_\infty}{\partial x} + V_\infty \frac{\partial W_\infty}{\partial y} + W_\infty \frac{\partial W_\infty}{\partial z} \end{bmatrix} \\
&= 0, \text{ since } \frac{\partial U_\infty}{\partial x} = \frac{\partial U_\infty}{\partial y} = \frac{\partial V_\infty}{\partial x} = \frac{\partial V_\infty}{\partial y} = 0 \text{ and } W_\infty = 0.
\end{aligned}$$

$$
\begin{aligned}
\text{T2.2} &= \mathbf{v_l} \cdot (\mathbf{V}_\infty \cdot \nabla)\mathbf{v_l} \\
&= \mathbf{v_l} \cdot \begin{bmatrix} U_\infty \frac{\partial u_l}{\partial x} + V_\infty \frac{\partial u_l}{\partial y} + W_\infty \frac{\partial u_l}{\partial z} \\ U_\infty \frac{\partial v_l}{\partial x} + V_\infty \frac{\partial v_l}{\partial y} + W_\infty \frac{\partial v_l}{\partial z} \\ U_\infty \frac{\partial w_l}{\partial x} + V_\infty \frac{\partial w_l}{\partial y} + W_\infty \frac{\partial w_l}{\partial z} \end{bmatrix} \\
&= U_\infty (u_l \frac{\partial u_l}{\partial x} + v_l \frac{\partial v_l}{\partial x} + w_l \frac{\partial w_l}{\partial x}) + V_\infty (u_l \frac{\partial u_l}{\partial y} + v_l \frac{\partial v_l}{\partial y} + w_l \frac{\partial w_l}{\partial y}) \\
&= U_\infty \frac{\partial E_l}{\partial dx} + V_\infty \frac{\partial E_l}{\partial dy}
\end{aligned}
$$

$$
\begin{aligned}
\text{T2.3} &= \mathbf{v_l} \cdot (\mathbf{V}_\infty \cdot \nabla)\mathbf{v'} \\
&= \mathbf{v_l} \cdot \begin{bmatrix} U_\infty \frac{\partial u'}{\partial x} + V_\infty \frac{\partial u'}{\partial y} + W_\infty \frac{\partial u'}{\partial z} \\ U_\infty \frac{\partial v'}{\partial x} + V_\infty \frac{\partial v'}{\partial y} + W_\infty \frac{\partial v'}{\partial z} \\ U_\infty \frac{\partial w'}{\partial x} + V_\infty \frac{\partial w'}{\partial y} + W_\infty \frac{\partial w' l}{\partial z} \end{bmatrix} \\
&= \mathbf{v_l} \cdot \begin{bmatrix} \frac{\partial (u' U_\infty)}{\partial x} + \frac{\partial (u' V_\infty)}{\partial y} + \frac{\partial (u' W_\infty)}{\partial z} \\ \frac{\partial (v' U_\infty)}{\partial x} + \frac{\partial (v' V_\infty)}{\partial y} + \frac{\partial (v' W_\infty)}{\partial z} \\ \frac{\partial (w' U_\infty)}{\partial x} + \frac{\partial (w' V_\infty)}{\partial y} + \frac{\partial (w' W_\infty)}{\partial z} \end{bmatrix} \text{ (using the continuity equation)} \\
&= 0, \text{ since } \frac{\partial}{\partial x_i} \widehat{a' b_\infty} = 0.
\end{aligned}
$$

$$
\begin{aligned}
\text{T2.4} &= \mathbf{v_l} \cdot (\mathbf{v_l} \cdot \nabla)\mathbf{V}_\infty \\
&= \mathbf{v_l} \cdot \begin{bmatrix} u_l \frac{\partial U_\infty}{\partial x} + v_l \frac{\partial U_\infty}{\partial y} + w_l \frac{\partial U_\infty}{\partial z} \\ u_l \frac{\partial V_\infty}{\partial x} + v_l \frac{\partial V_\infty}{\partial y} + w_l \frac{\partial V_\infty}{\partial z} \\ u_l \frac{\partial W_\infty}{\partial x} + v_l \frac{\partial W_\infty}{\partial y} + w_l \frac{\partial W_\infty}{\partial z} \end{bmatrix} \\
&= u_l w_l \frac{\partial U_\infty}{\partial z} + v_l w_l \frac{\partial V_\infty}{\partial z} \\
&\text{since } \frac{\partial U_\infty}{\partial x} = \frac{\partial U_\infty}{\partial y} = \frac{\partial V_\infty}{\partial x} = \frac{\partial V_\infty}{\partial y} = 0 \text{ and } W_\infty = 0.
\end{aligned}
$$

$$
\begin{aligned}
\text{T2.5} &= \mathbf{v_l} \cdot (\mathbf{v_l} \cdot \nabla)\mathbf{v_l} \\
&= \mathbf{v_l} \cdot \begin{bmatrix} u_l \frac{\partial u_l}{\partial x} + v_l \frac{\partial u_l}{\partial y} + w_l \frac{\partial u_l}{\partial z} \\ u_l \frac{\partial v_l}{\partial x} + v_l \frac{\partial v_l}{\partial y} + w_l \frac{\partial v_l}{\partial z} \\ u_l \frac{\partial w_l}{\partial x} + v_l \frac{\partial w_l}{\partial y} + w_l \frac{\partial w_l}{\partial z} \end{bmatrix} \\
&= u_l u_l \frac{\partial u_l}{\partial x} + u_l v_l \frac{\partial u_l}{\partial y} + u_l w_l \frac{\partial u_l}{\partial z} \\
&\quad + v_l u_l \frac{\partial v_l}{\partial x} + v_l v_l \frac{\partial v_l}{\partial y} + v_l w_l \frac{\partial v_l}{\partial z} \\
&\quad + w_l u_l \frac{\partial w_l}{\partial x} + w_l v_l \frac{\partial w_l}{\partial y} + w_l w_l \frac{\partial w_l}{\partial z} \\
&= u_l \frac{\partial E_l}{\partial x} + v_l \frac{\partial E_l}{\partial y} + w_l \frac{\partial E_l}{\partial z} \\
&= u_{j,l} \frac{\partial E_l}{\partial x_i}
\end{aligned}
$$

$$
\begin{aligned}
\text{T2.6} &= \mathbf{v_l} \cdot (\mathbf{v_l} \cdot \nabla)\mathbf{v'} \\
&= \mathbf{v_l} \cdot \begin{bmatrix} u_l \frac{\partial u'}{\partial x} + v_l \frac{\partial u'}{\partial y} + w_l \frac{\partial u'}{\partial z} \\ u_l \frac{\partial v'}{\partial x} + v_l \frac{\partial v'}{\partial y} + w_l \frac{\partial v'}{\partial z} \\ u_l \frac{\partial w'}{\partial x} + v_l \frac{\partial w'}{\partial y} + w_l \frac{\partial w'}{\partial z} \end{bmatrix} \\
&= \mathbf{v_l} \cdot \begin{bmatrix} \frac{\partial u' u_l}{\partial x} + \frac{\partial u' v_l}{\partial y} + \frac{\partial u' w_l}{\partial z} \\ \frac{\partial v' u_l}{\partial x} + \frac{\partial v' v_l}{\partial y} + \frac{\partial v' w_l}{\partial z} \\ \frac{\partial w' u_l}{\partial x} + \frac{\partial w' v_l}{\partial y} + \frac{\partial w' w_l}{\partial z} \end{bmatrix} \text{ (using the continuity equation)} \\
&= 0, \text{ since } \widehat{a'b_l} = 0.
\end{aligned}
$$

$$
\begin{aligned}
\text{T2.7} &= \mathbf{v_l} \cdot (\mathbf{v'} \cdot \nabla)\mathbf{V}_\infty \\
&= \mathbf{v_l} \cdot \begin{bmatrix} u' \frac{\partial U_\infty}{\partial x} + v' \frac{\partial U_\infty}{\partial y} + w' \frac{\partial U_\infty}{\partial z} \\ u' \frac{\partial V_\infty}{\partial x} + v' \frac{\partial V_\infty}{\partial y} + w' \frac{\partial V_\infty}{\partial z} \\ u' \frac{\partial W_\infty}{\partial x} + v' \frac{\partial W_\infty}{\partial y} + w' \frac{\partial W_\infty}{\partial z} \end{bmatrix} \\
&= 0 \text{ since } \widehat{a_l b'} = 0.
\end{aligned}
$$

$$
\begin{aligned}
\text{T2.8} &= \mathbf{v_l} \cdot (\mathbf{v'} \cdot \nabla)\mathbf{v_l} \\
&= \mathbf{v_l} \cdot \begin{bmatrix} u' \frac{\partial u_l}{\partial x} + v' \frac{\partial u_l}{\partial y} + w' \frac{\partial u_l}{\partial z} \\ u' \frac{\partial v_l}{\partial x} + v' \frac{\partial v_l}{\partial y} + w' \frac{\partial v_l}{\partial z} \\ u' \frac{\partial w_l}{\partial x} + v' \frac{\partial w_l}{\partial y} + w' \frac{\partial w_l}{\partial z} \end{bmatrix} \\
&= 0 \text{ since } \widehat{a_l b'} = 0.
\end{aligned}
$$

$$
\begin{aligned}
\text{T2.9} &= \mathbf{v_l} \cdot (\mathbf{v}' \cdot \nabla)\mathbf{v}' \\
&= \mathbf{v_l} \cdot \begin{bmatrix} u'\frac{\partial u'}{\partial x} + v'\frac{\partial u'}{\partial y} + w'\frac{\partial u'}{\partial z} \\ u'\frac{\partial v'}{\partial x} + v'\frac{\partial v'}{\partial y} + w'\frac{\partial v'}{\partial z} \\ u'\frac{\partial w'}{\partial x} + v'\frac{\partial w'}{\partial y} + w'\frac{\partial w'}{\partial z} \end{bmatrix} \\
&= \mathbf{v_l} \cdot \begin{bmatrix} \frac{\partial u'u'}{\partial x} + \frac{\partial v'u'}{\partial y} + \frac{\partial w'u'}{\partial z} \\ \frac{\partial u'v'}{\partial x} + \frac{\partial v'v'}{\partial y} + \frac{\partial w'v'}{\partial z} \\ \frac{\partial u'w'}{\partial x} + \frac{\partial v'w'}{\partial y} + \frac{\partial w'w'}{\partial z} \end{bmatrix} \quad \text{(using the continuity equation)} \\
&= u_{j,l}\frac{\partial}{\partial x_i}u_i u_j
\end{aligned}
$$

Non-zero contributions arise from T2.2, T2.4, T2.5 and T2.9. Combining these terms gives:

$$
\text{T2} = U_\infty \frac{\partial E_l}{\partial dx} + V_\infty \frac{\partial E_l}{\partial dy} + u_l w_l \frac{\partial U_\infty}{\partial z} + v_l w_l \frac{\partial V_\infty}{\partial z} + u_{j,l}\frac{\partial E_l}{\partial x_i} + u_{j,l}\frac{\partial}{\partial x_i}u_i u_j
$$

Term 3

$$
\begin{aligned}
\text{T3} &= \mathbf{v_l} \cdot \frac{1}{\rho}\nabla P \\
&= \mathbf{v_l} \cdot \frac{1}{\rho}(\nabla P_\infty + \nabla P_l + \nabla P' \\
&= \frac{1}{\rho}\mathbf{v_l} \cdot \begin{bmatrix} \frac{\partial P_\infty}{\partial x} + \frac{\partial P_l}{\partial x} + \frac{\partial P'}{\partial x} \\ \frac{\partial P_\infty}{\partial y} + \frac{\partial P_l}{\partial y} + \frac{\partial P'}{\partial y} \\ \frac{\partial P_\infty}{\partial z} + \frac{\partial P_l}{\partial z} + \frac{\partial P'}{\partial z} \end{bmatrix} \\
&= \frac{1}{\rho}\begin{bmatrix} \frac{\partial (P_\infty u_l)}{\partial x} + \frac{\partial (P_l u_l)}{\partial x} + \frac{\partial (P' u_l)}{\partial x} \\ \frac{\partial (P_\infty v_l)}{\partial y} + \frac{\partial (P_l v_l)}{\partial y} + \frac{\partial (P' v_l)}{\partial y} \\ \frac{\partial (P_\infty w_l)}{\partial z} + \frac{\partial (P_l w_l)}{\partial z} + \frac{\partial (P' w_l)}{\partial z} \end{bmatrix} \quad \text{(using the continuity equation)} \\
&= \frac{1}{\rho}\frac{\partial P_l u_l}{\partial x} + \frac{\partial P_l v_l}{\partial y} + \frac{\partial P_l w_l}{\partial z} \\
&= \frac{1}{\rho}\frac{\partial}{\partial x_i}(P_l u_{i,l})
\end{aligned} \tag{B.6}
$$

Term 4

$$
\begin{aligned}
\text{T4} &= g\frac{\theta_v}{\theta_{v,0}}\mathbf{k} \\
&= g\frac{w_l\theta_{v,\infty} + \theta_{v,l} + \theta'_v}{\theta_{v,0}}\mathbf{k} \\
&= \frac{g}{\theta_{v,0}} w_l \theta_{v,l}
\end{aligned}
$$

Term 5

$$
\begin{aligned}
\mathrm{T5} &= \mathbf{v_l} \cdot 2\mathbf{\Omega_e} \times \mathbf{V} \\
&\quad \text{where } \mathbf{\Omega_e} = (0, \Omega_{e,1}, \Omega_{e,2}) \\
&= u_l(\Omega_{e,1} W - \Omega_{e,2} V) + v_l(\Omega_{e,2} U) - w_l(\Omega_{1,e} U) \\
&= u_l(\Omega_{e,1} w_l - \Omega_{e,2} v_l) + v_l(\Omega_{e,2} u_l) - w_l(\Omega_{1,e} u_l) \\
&\quad \text{since } \widehat{a' b_l} = 0 \text{ and } \widehat{a}_l = 0 \\
&= u_l w_l \Omega_{e,1} - u_l w_l \Omega_{e,1} + u_l v_l \Omega_{e,2} - u_l v_l \Omega_{e,2} \\
&= 0
\end{aligned} \tag{B.7}
$$

Combining terms 1–5, the budget equation is:

$$
\begin{aligned}
\frac{\partial \hat{E}_l}{\partial t} = &-U_\infty \frac{\partial \hat{E}_l}{\partial x} - V_\infty \frac{\partial \hat{E}_l}{\partial y} \underbrace{- \widehat{u_l w_l} \frac{dU_\infty}{dz} - \widehat{v_l w_l} \frac{dV_\infty}{dz}}_{\text{shear}} + \underbrace{\frac{g}{\theta_{v,\infty}} \widehat{\theta_{v,l} w_l}}_{\text{buoyancy}} \\
&\underbrace{- \widehat{u_{j,l}} \frac{\partial \hat{E}_l}{\partial x_j}}_{\text{cell-scale advection}} - \underbrace{\frac{1}{\rho} \frac{\partial}{\partial x_j} \widehat{p_l u_{j,l}}}_{\text{pressure-wind balance}} \underbrace{- u_{k,l} \frac{\partial}{\partial x_j} (\widehat{u'_k u'_j})_l}_{\text{interscale transfer}}
\end{aligned} \tag{B.8}
$$

Resolving nonstationary spectral information in wind speed time series using the Hilbert-Huang transform

---

Authors: Claire Vincent, Gregor Giebel, Pierre Pinson and Henrik Madsen
Journal of Applied Meteorology and Climatology, 2010(49) pages 253–267

Wind fluctuations over the North Sea

Authors: Claire L Vincent, Pierre Pinson and Gregor Giebel
International Journal of Climatology 2010, DOI: 10.1002/joc.2175

Idealised WRF simulations of open cellular convection over the sea

Claire L Vincent, Andrea N Hahmann and Mark C Kelly
Submitted to Boundary Layer Meteorology

Risø DTU is the National Laboratory for Sustainable Energy. Our research focuses on development of energy technologies and systems with minimal effect on climate, and contributes to innovation, education and policy. Risø has large experimental facilities and interdisciplinary research environments, and includes the national centre for nuclear technologies.

Risø DTU
National Laboratory for Sustainable Energy
Technical University of Denmark

Frederiksborgvej 399
PO Box 49
DK-4000 Roskilde
Denmark
Phone +45 4677 4677
Fax +45 4677 5688

www.risoe.dtu.dk